Aalto University
School of Business

Designing context-appropriate interactions for virtual reality

Investigating methods of collaboration between immersed and external users in asymmetric virtual reality.

Master's Thesis
Markus Kirjonen
Aalto University School of Business
International Design Business Management
Fall 2020

Acknowledgements

The experiment described in this thesis was made possible by my project supervisor, Mika P. Nieminen. Not only did he provide the initial research direction and funding necessary to conduct our user studies, but also continued optimism, guidance in conducting user studies, and valuable insights into our experiment design, for which I am deeply grateful.

I would also like to thank my thesis supervisor, Miikka J. Lehtonen, for his enthusiasm and guidance throughout the entirety of this thesis process. His support, both academic and personal, has been highly appreciated.

A short paper describing this experiment setup was published prior to performing the experiment (Nieminen & Kirjonen, 2020).

Aalto University
School of Business

Aalto University, P.O. BOX 11000, 00076 AALTO
www.aalto.fi

Abstract of master's thesis

Author Markus Kirjonen

Title of thesis Designing context-appropriate interactions for virtual reality

Degree MSc. In Economics and Business Administration

Degree International Design Business Management

Thesis advisor(s) Miikka J. Lehtonen

Year of approval 2020 **Number of pages** 60 **Language** English

Abstract

This thesis explores the concept of presence in virtual reality. A high degree of presence has been shown to be essential in eliciting realistic emotions and responses in virtual reality, which in turn makes virtual reality a useful tool for studying human behavior in contexts such as design, psychology, user research, and more. For instance, researchers have been using virtual reality as a cost-effective method of prototyping products and environments in cases where immersing the user in a similar context in the real world would be cost-prohibitive, dangerous, or simply unpractical or impossible. Examples of this include allowing would-be patients to experience a yet-to-be-designed hospital room, co-designing device interfaces in collaboration with one's target audience, designing the interiors of stores and monitoring which items catch people's attention, or studying how people navigate through an airport terminal that would otherwise be off-limits to designers. By observing how people respond to these virtual product and environments, practitioners can better understand how their design choices will play out in the real world before any physical production begins, allowing them to make changes that would otherwise be costly later on in the design process.

Unfortunately, presence is also highly susceptible to being broken - especially when interacting with the real world. In fact, research has shown that the most common cause for breaks in presence while in virtual reality is interference from the outside world. This presents a challenge for researchers and practitioners hoping to use virtual reality with their users, as their experiment designs often involve immersed users interacting with external users. For instance, in the context of a design study, an external moderator may want to ask the immersed user to perform certain actions or to describe their subjective experience of the contents of the virtual environment. These kinds of interactions, while necessary components of such studies, are highly likely to interrupt the immersed user's sense of presence, which in turn reduces the virtual environment's ability to provoke realistic behavior, thus reducing the effectiveness of the entire experiment. Fortunately, it seems that this may be avoidable; several researchers have presented preliminary examples of how integrating interactions into the context of a virtual environment can help prevent these breaks in presence. This thesis aims to expand on these studies.

Here, we discuss contexts where users in the real world would want to interact with users immersed in virtual reality, and how such interactions might be designed to avoid disturbing the immersed user's sense of presence. We describe an experiment we conducted in order to test out various methods of communication and collaboration between immersed and external users to that end. Based on our results, we make design suggestions for researchers hoping to use virtual reality in their own experiments, with a focus on domains such as user experience research, design work, and psychology.

Keywords design, virtual reality, asymmetric virtual reality, collaboration

Table of Contents

List of Tables and Figures

Tables

Figures

1. Introduction

This thesis discusses the concept of presence, the perception of being physically present in a virtual environment, in the context of multi-user virtual reality. We investigate communication methods and interactions between users in asymmetric virtual reality, with the goal of increasing the effectiveness of virtual reality in domains such as collaborative problem-solving, co-design, and research on human behavior.

Presence is a key element of virtual reality. It has often been described as a user's subjective psychological response to a virtual environment, or more succinctly, the feeling of "being there." (Heeter, 1992; Lombard & Ditton, 2006; Slater, 2003). It has been shown that when a user experiences a high degree of presence, they respond to the virtual environment as they would to the real world, exhibiting realistic reactions, behaviors, and emotions (Slater, Khanna, Mortensen & Yu, 2009; Diemer, Alpers, Peperkorn, Shiban & Mühlberger, 2015). Further, the more present a user is, the more realistic their responses are (Kober & Neuper, 2012).

This ability to elicit realistic responses makes virtual reality a promising tool beyond mere entertainment; indeed, it has already seen adoption amongst user experience and usability designers (Brade et al., 2017; Rebelo, Noriega, Duarte & Soares, 2012; Tiainen & Jouppila, 2019), psychiatrists (Botella, Fernández-Álvarez, Guillén, García-Palacios & Baños, 2017; Ling, Nefs, Morina, Heynderickx & Brinkman, 2014; Price & Anderson, 2007; Koller, Schafer, Lochner & Meixner, 2019), psychologists (Diemer, Alpers, Peperkorn, Shiban & Mühlberger, 2015; Slater et al., 2006; Yee & Bailenson, 2007), and more. With the click of a mouse, a researcher can place a user in a simulation of a wide variety of real-world scenarios and environments, many of which would otherwise be too dangerous, cost-prohibitive, or simply impractical to conduct research in. This offers a potentially significant improvement over more common methods of immersing a user in a different context, such as design artifacts or narratives. These potential use cases are discussed in section 2.5. When conducting virtual reality sessions in these contexts, it is common for one user, the subject, to be equipped with a virtual reality headset and immersed in the virtual environment, while another user, often the moderator, remains outside of it. This setup is referred to as asymmetric virtual reality and is the basis for our experiment. Asymmetric virtual reality is explored in section 2.3.

The potential of virtual reality in a research context rests upon its ability to elicit realistic responses from immersed users – and as long as the users maintain a high degree of presence, this will be the case. Unfortunately, presence in virtual reality is not a given (Bowman & McMahan, 2007; Cummings & Bailenson, 2015). There are a wide variety of factors that affect how present a user feels at a given moment, and it is still an active area of research. Design choices within the virtual environment are highly relevant in establishing a strong sense of presence (Usoh, Catena, Arman & Slater, 2000; Schuemie, Straaten, Krijn & van der Mast, 2001; Bowman & McMahan, 2007). Other commonly described factors relate to the technology itself, with features such as the level of tracking, field of view, and update rate often being cited (Cummings & Bailenson, 2015; Weech, Kenny & Barnett-Cowan, 2019). In addition, we also see factors that relate to the individual user; as presence is a subjective, context-based phenomenon, it is dependent on the user's state of mind going into the virtual environment, as well as their disposition to immersive experiences and personality traits (Samana, Wallach & Safir, 2009). Factors that contribute to and decrease presence are described in section 2.2. Depending on these factors, a user's sense of presence can fluctuate throughout the duration of a virtual reality session and, in the most extreme cases, may even break entirely – referred to as a break in presence (Slater & Steed, 2000). In their description of the phenomenon of breaks in presence, Slater & Steed (2000) classified reasons for such occurrences, finding two primary categories: 1) external causes, where sensory information from the external world intrudes on or contradicts what is happening in the virtual world, and 2) internal causes, where something is "wrong" in the virtual world, such as unrealistic physics or visuals. Breaks in presence are discussed in section 2.2. and methods of measuring presence in section 2.4.

Distractions from the external world, *external causes* under Slater & Steed's (2000) classification, provide some of the most common reasons for breaks in presence. The two top reasons for breaks in presence are hearing noises from the real world, such as a person speaking, and experiencing external touches or forces. However, this disruption to a user's sense of presence seems to occur in any situation where the user interacts with, or becomes aware of, the external world (Liszio & Masuch, 2016). While this is a challenge that any virtual reality setup has to cope with, it is especially difficult when dealing with asymmetric virtual reality, which is all about interactions between immersed users in virtual reality and external users in the real world.

Several authors have investigated the concept of presence in asymmetric virtual reality contexts and how it can be maintained, despite the challenges that come along with attempting to bridge the real and virtual worlds. Zenner et al. (2018) found that externally caused breaks in presence are not limited to audio and physical forces; rather, all digital information (texts, calls, calendar events, notifications) from the external world can lead to a break in presence – if not adapted to the virtual environment's context. This last part, adaption, seems to be key. They found that by integrating notifications into the immersed user's virtual environment in the form of messages on television screens, the external user was able to communicate with the immersed user without disrupting their sense of presence. This aligns with earlier research into presence in traditional video games; as Cairns et al. point out, only interactions that don't happen within the context of a virtual environment are potential disruptors of presence (Cairns, Cox, Day, Martin & Perryman, 2013). Koller, Schafer, Lochner & Meixner (2019) further explored this concept of integration in the context of therapy. They developed a system that allowed a therapist, the external user, to take control of an audience member during a public speaking simulation, using the avatar to communicate with the immersed user. They found that this kind of direct verbal interaction between the external and immersed users enhanced the immersed user's experience and increased the efficiency of the therapy process.

This thesis builds on these findings. As Zenner et al. (2018) and Koller, Schafer, Lochner & Meixner (2019) have shown, it is possible to enable interactions between external and immersed users in an asymmetric virtual reality context as long as a certain level of integration into the virtual environment is achieved. However, this brings up an important issue - while these authors show that integrating interactions into the virtual environment was conducive to presence, and that attempting to communicate through non-integrated means often was not, the question of how different levels of integration affect the immersed user's sense of presence remains. That is, can higher levels of presence be achieved through higher degrees of integration? And going further, how does the dynamic between the immersed and external user change based on the degree of integration? The underlying assumption here being that the nature of the integration will influence how the immersed user perceives and feels about the external user – after all, receiving a notification on a television screen, as seen in Zenner et al. (2018), is presumably very different than speaking to Koller, Schafer, Lochner & Meixner (2019)'s embodied audience members.

Based on these questions, we can establish our two primary research questions:

RQ1: In asymmetric virtual reality, how do varying levels of integrating interactions into a virtual environment affect the immersed user's sense of presence?

RQ2: How does the dynamic between the immersed and external user change based on the degree of interaction integration in asymmetric virtual reality?

In order to investigate these questions, we developed a virtual environment featuring multiple methods of integrating interactions into the environment (detailed in section 3.4). The environment features a series of tasks that the immersed user must solve with the assistance of an external user. During each task, one of our five methods of interaction was randomly selected to facilitate the interaction between the two users. Some of these methods of interaction were designed to be highly integrated into the environment, while others were intentionally unintegrated. A full list of interaction methods is available in section 3, Table 1.

Fifteen users were selected to play through the virtual environment. Our selection criteria and experiment procedure are described in section 3.4.3. During each of these user sessions, we used a combination of short questionnaires, user comments made during the session, and a semi-structured post-session interview to understand how the user's sense of presence fluctuated throughout the experience, as well as how they perceived the external user's role in the environment. As we came to find, the user interviews provided the most insight into how the users experienced the virtual environment and the dynamic between them and the external user. As such, a significant portion of section 4, Findings, is spent exploring themes that came up during these interviews.

Going into the experiment, our initial hypothesis was that the more integrated interactions were into the environment, the higher the immersed user's sense of presence would be. It soon became clear that a significantly more nuanced understanding of how integrating interactions in asymmetric virtual reality affected immersed users was necessary; the relationship between how interaction integration and how the immersed user's sense of presence was affected was not that straightforward. Instead, we found that the degree, and more importantly, style of integration changed the component of presence being acted upon (social presence versus spatial presence). Further, we found that how the immersed user perceived the role of the external user changed based on the style of integration. Additionally, several other important factors, such as how much control the external user has over the interactions, were affected. These findings are explored in depth in section 4, and their design implications are explored in sections 5 and 6.

The rest of the thesis is structured as follows: we begin with a review of Literature regarding presence in virtual reality, in both standard and asymmetric contexts. Here, special attention is paid to previous attempts to integrate interactions into the virtual environment during asymmetric sessions. Following that, we describe the experiment setup and procedure in the Methodology. Here we describe the environment we built and the components that make it up. From there, we move onto the Findings, where we explore the data that was generated during user sessions and post-session interviews. The themes that we develop in that section are further explored in the Discussion, where we make design recommendations based on what we found during our sessions. Finally, in the Conclusion, we summarize the purpose of the experiment, our experiment design, our findings, and our key takeaways.

2. Literature review

In this thesis, we build upon research into presence in virtual reality in both standard and asymmetric contexts. Here, standard refers to traditional virtual reality setups, featuring either a single user or multiple remote users with identical roles and abilities. Asymmetric, on the other hand, refers to virtual reality setups where there is an asymmetry in the immersion of users (e.g. one user is immersed in virtual reality while the other remains in the real world) as well as the user's abilities to modify and see the virtual environment. The former, standard virtual reality setups, have been much more heavily researched. As such, it forms the bulk of research on the topic, with many papers on asymmetric virtual reality setups being limited to the latter part of the past decade. Much of the research into presence in standard virtual reality contexts is directly applicable to asymmetric contexts. The key point of contention between these two areas of research rests in asymmetric interactions. In early papers on presence, such as Slater & Steed (2000)'s classification of causes of breaks in presence, external influences from the real world are often treated as undesired interruptions. In asymmetric contexts, external influences are often an essential component. Beyond this deviation, the general conceptualizations of presence and methods of measuring it remain consistent.

This section outlines many of the important concepts related to presence in virtual reality. After briefly expanding upon the previously provided definition of presence, we take a look at factors that can influence presence and explore the concept of breaks in presence. Following that, we shift to asymmetric virtual reality and how authors have handled presence in that context. Then, we look at methods of measuring presence. Finally, we finish with examples of virtual reality, both standard and asymmetric, being used in a variety of non-entertainment use cases and suggest how the research questions in this thesis may be relevant to such situations.

2.1. Presence

In the beginning of this thesis, presence was described as a user's subjective psychological response to a virtual environment (Heeter, 1992; Lombard & Ditton, 2006; Slater, 2003). A key concept here, as we will see when we explore methods of measuring presence, is that it is subjective and psychological; it occurs in the mind of the user.

Presence is often used in conjunction with the term “immersion.” In fact, these terms are occasionally used interchangeably, despite representing distinct concepts, which can lead to confusion (Bowman & McMahan, 2007). Several authors have provided differentiations between the two concepts, with a popular one being offered by prominent virtual reality researcher Mel Slater. Slater emphasized that presence is an individual, subjective response on the part of the user, while immersion is an objective description of the virtual reality setup’s ability to “immerse the user.” (Slater 1999; Slater, 2003). Under this definition, immersion is objective and easily measurable, defined by hardware and software factors such as resolution, quality of tracking, and field of view. Presence, on the other hand, is context dependent and can vary across identical virtual reality systems, or even for one user in the same system depending on their state of mind going into the experience (Bowman & McMahan, 2007).

Researchers have distinguished between different components of presence, three of which have been prominently featured in literature: social presence, the sense of being physically located amongst others, spatial presence, the sense of being physically located in the virtual environment, and self-presence, the sense of experiencing one’s virtual avatar as being oneself (Heeter, 1992; Lee 2004; Tamborini & Bowman, 2010).

2.2. Factors that influence presence

There are a wide variety of factors that influence a user’s sense of presence.

Some commonly described factors relate to the technology itself. Features of technology – especially the tracking level, field of view, update rate, and stereoscopy, have been shown to have a significant effect on presence (Cummings & Bailenson, 2015). Slater, Khanna, Mortensen & Yu (2009) suggest that these features allow for a low level of latency between sensory data and proprioception; i.e. the location that we see our arm in matches the location that we *feel* our arm in. Less impactful but still influential features include image quality and resolution (Cummings & Bailenson, 2015). Limits of technology can also lead to breaks in presence. Cybersickness, which can be brought on by factors such as latency in tracking or low refresh rates, is a well-known cause of breaks in presence (Weech, Kenny & Barnett-Cowan, 2019). Further, awareness of the virtual reality equipment itself, whether the headset, battery, or controllers, can also lead to a decreased sense of presence (Riches, Elghany, Garety, Rus-Calafell & Valmaggia, 2019). As virtual reality hardware continues to improve, it is likely that we will see the technological factors relating to increased presence improve, while the technology-related drops and breaks in presence become less common.

Design choices within the virtual environment are also highly relevant in establishing a strong sense of presence. A user's sense of presence can be positively influenced by paying attention to the virtual environment's reactions to the actions of the user. Examples of this include realistic lighting and shadows (Bowman & McMahan, 2007), having objects that represent living entities behave realistically and respond to both the user and environment around them (Slater, Khanna, Mortensen & Yu, 2009), and providing the user with a high degree of ability to participate in and modify the virtual environment around them (Schuemie, Straaten, Krijn & van der Mast, 2001). Choices such as locomotion, how the user moves around the space, are also important. It has been shown that users experience a higher level of presence when locomotion is performed by tracking the movement of the user in the real world, rather than using a button to slide forward (Usoh, Catena, Arman & Slater, 2000). Similarly, it has been shown that techniques such as hand-tracking, which remove the need for a controller or conventional input device, lead to increased levels of presence (Jeong, Kim, Kim, Lee & Kim, 2019; Han & Kim, 2017). Design choices are also extremely relevant when it comes to asymmetric virtual reality contexts – as discussed previously, how the designer of the virtual environment chooses to integrate interactions between the immersed and external users will determine whether the interactions enable or hinder presence (Koller, Schafer, Lochner & Meixner, 2019; Zenner et al., 2018).

And of course, as emphasized earlier, presence is a subjective, context-based phenomenon, so some factors will naturally relate to the individual user in question. Individual traits that have been found to influence presence include imagination, empathy, level of anxiety, cognitive style, dissociative tendencies, and more (Samana, Wallach & Safir, 2009). The user's situational interest is also an important factor. Innate interest in the virtual environment's content can contribute to a stronger sense of presence (Diemer, Alpers, Peperkorn, Shiban & Mühlberger, 2015; Lessiter, Freeman, Keogh & Davidoff, 2001). Similarly, a user's general state of mind at the time of the experience and recent history are influential too (Bowman & McMahan, 2007).

2.3. Asymmetric virtual reality

Many virtual reality environments allow for multiple users to co-exist in a shared virtual space. In these environments, users can interact, collaborate, and communicate with each other. Users don't always share the same level of immersion in multi-user virtual environments. For instance, you may have a user wearing a virtual reality headset interacting and communicating with non-immersed participants. This is referred to as asymmetric virtual reality.

Several popular virtual reality games have used this asymmetry in immersion as a part of the core gameplay. Examples include Keep Talking and Nobody Explodes, Ruckus Ridge VR Party, Black Hat Cooperative, and Eye in the Sky. Researchers have also experimented with asymmetric interfaces. ShareVR (Gugenheimer, Stemasov, Frommel & Rukzio, 2017) and RoleVR (Lee, Kim & Kim, 2019) are examples of this, with the respective authors proposing asymmetric interactions designed to provide satisfying experiences and higher levels of presence for both immersed and non-immersed users. Asymmetry has also been explored in the context of design; for instance, Sugiura et al. (2018) developed a system for asymmetric architectural design. In their setup, the immersed user sees the interior of the building at proper scale, while the external user manipulates it.

Even though the term may not explicitly be used, asymmetric virtual reality setups are quite common outside of entertainment. When conducting a virtual reality session, whether one is engaged in collaborative design, psychiatry, or numerous other academic or professional contexts, it is common to see an immersed user inside the virtual environment while an external user acts as a moderator, providing instructions and modifying the environment as needed. Horst, Dörner & Peter (2018) identified a rich set of features that these moderators may use, categorizing them into four categories: 1) View-related features, which allow the moderator to observe the virtual scene from multiple perspectives; 2) Manipulation features, which allow the moderator to control, add, and delete objects in the scene; 3) Meta-part features, which allow the moderator to control the simulation; and 4) Monitoring features, which allow the moderator to gather affective computing data such as the immersed user's heart rate or skin conductivity.

When immersed and external users interact in asymmetric virtual reality contexts, there is a potential to disrupt the immersed user's sense of presence by calling their attention to the external world. As we have seen, this is a challenge facing asymmetric setups. Several authors have investigated presence in asymmetric virtual reality. Gugenheimer, Stemasov, Frommel & Rukzio (2017) initiated this by proposing ShareVR, a proof-of-concept prototype that allowed immersed and non-immersed users to interact with increased presence for both parties. Building on their work, Lee, Kim & Kim (2019) presented RoleVR, a system designed to provide both immersed and non-immersed users with high levels of presence when collaborating in an asymmetric virtual environment. At the same time, Jeong, Kim, Kim, Lee & Kim (2019) proposed an asymmetric virtual reality interface also designed to provide both users with a sense of presence during interactions. Several subsequent authors have focused on the immersed user's sense of presence, rather than that of both users. Zenner et al. (2018) proposed a framework for integrating mobile notifications in virtual reality while preserving the immersed user's immersion and suspension of disbelief. Their goal was to allow non-immersed users to get in contact with the immersed user without causing a break in presence. Koller, Schafer, Lochner & Meixner (2019) developed a system that allowed therapists to take control of aspects of a virtual environment that their patient was immersed in, enabling non-disruptive communication between the two parties. Finally, Liszio & Masuch (2016) proposed a method for integrating communication and interaction between immersed and external players into virtual reality game mechanics, allowing for interaction without breaking the player's presence.

2.4. Measuring presence

As we have seen, presence is an important concept in virtual reality as it enables realistic user responses to the virtual environment. We have taken a look at many of the primary factors that increase and decrease presence and have explored the concept of breaks in presence. However, before we can integrate these concepts into our experiment, we need to establish a method of measuring how changes to our virtual environment affect our user's sense of presence. As we will see, this is not a straightforward task.

Researchers have suggested numerous ways to measure presence, often distinguished into subjective and objective methods. To help deal with the wide variety of different approaches, Van Barren & IJsselsteijn provide a detailed list of methods for measuring presence and relevant examples for each (van Baren & IJsselsteijn, 2004).

Objective methods of measuring presence include recording physiological responses (heart rate, skin temperature and conductance, neuroimaging) and behavioral responses (studying social and reflexive responses, posture, responses to conflicting cues). While objective methods are appealing to researchers, many authors have emphasized the limitations of studying presence through such procedures. Cummings & Bailenson (2015) argue that it is difficult to discern what specific stimuli in a virtual environment is causing the physiological or behavioral response. As one can imagine, this problem becomes more significant as the virtual environment's complexity increases. Lombard et al. provide additional arguments, pointing out that it is difficult to use these objective measures outside of high-action scenarios (Lombard, Bolmarcich & Weinstein, 2009). A subject's heart rate may increase during a virtual shootout – but what about while sitting calmly on a virtual airplane? The user may very well be experiencing a high degree of presence in both situations, despite the latter lacking detectable objective responses such as an increase in heart rate or brain activity as a result of the relatively peaceful nature of the experience. Due to these limitations, many researchers have concluded that for now, objective measures are a good addition to asking users to describe their subjective experience, not a replacement (Lombard, Bolmarcich & Weinstein, 2009; Cummings & Bailenson, 2015). Nevertheless, researchers are excited by the future potential of objective measures, especially when considering the limitations of subjective approaches - as explored next (Weech, Kenny & Barnett-Cowan, 2019).

When it comes to subjective methods of measuring presence, questionnaires are the most common method of measuring a user's sense of presence in virtual reality in general. These questionnaires are designed to be delivered after a virtual reality experience and have the user reflect on how they felt during the session. While questionnaires are the most common form of measuring presence, their post-experience nature has been criticized. As they are usually administered after the session, questionnaires rely on the user's memory of presence, rather than the experience of presence itself (Usoh, Catena, Arman & Slater, 2000; Schwind et al., 2017). Further, as Slater (2004) points out, by asking about one's sense of presence, you are bringing into question the phenomenon that you're supposed to be measuring, which may in turn affect the phenomenon itself. And finally, on a more practical note, administering a questionnaire requires the user to leave the virtual environment and remove the headset, which is both guaranteed to break their presence and takes up time during the user session (Schwind, Knierim, Haas & Henze, 2019). To help circumvent these issues while still relying on subjective means, some researchers have suggested self-reporting during the experience rather than afterwards. One method for doing this is integrating questionnaires into virtual reality. This concept was borrowed from presence research in desktop gaming (Frommel et al., 2015) and has been explored by e.g. Schwind et al. in two experiments (Schwind, Knierim, Chuang & Henze, 2017; Schwind et al., 2017). Interviews and post-session discussions are also examples of subjective methods of understanding how users experienced a virtual environment.

2.5. Virtual reality use-cases

To conclude our exploration of previous work, we'll take a look at a few examples of virtual reality being used in a variety of non-entertainment contexts where presence is important. This will help us better understand how our research questions may be of use to both current virtual reality designers as well as potential practitioners who are curious about how virtual reality might be of use in their own endeavors.

Usability testing and user experience design are design-related areas that have adopted virtual reality. As Brade et al. (2017) showed, virtual environments can be used as an alternative to real environments for user experience studies, provided a high level of presence is achieved. This allows for cost effective prototyping and co-designing of products, services, and environments (Rebelo, Noriega, Duarte, Soares, 2012). Examples of this include Tiainen & Jouppila's (2019) study on hospital design, where would-be patients described their subjective reactions to the space prior to construction and Mobach (2008)'s study which involved users in the design of pharmacies. Rebelo, Noriega, Duarte & Soares (2012) provide an overview of how virtual reality may be used for user testing in this context.

Virtual reality has also been gaining ground in psychology, as it allows us to study psychological and social dynamics in highly controlled, highly customizable settings (Diemer, Alpers, Peperkorn, Shiban & Mühlberger, 2015). An early example of using virtual reality to perform studies that would otherwise not be possible in real life involved Slater et al.'s (2006) repeat of the infamous Milgram obedience experiments in virtual reality, avoiding the ethical objections that have prevented repeats of the experiment in the past. In another example of virtual reality enabled experiments, Yee & Bailenson (2007) investigated the effect of self-representation on behavior. They found that people's behavior changed depending on the avatars that they embodied. When embodying a taller avatar, subjects acted with more confidence, and when provided with a more attractive avatar, they were more intimate with their disclosures. In a similar experiment, Peck, Seinfeld, Aglioti & Slater (2013) had users embody an avatar that represented a minority race in order to study racial bias reduction.

Similarly, virtual reality has proven itself in the field of psychiatric exposure therapy. Here, therapists can use virtual reality to deliver individualized and fully controlled phobic stimulus. This potentially offers a significant improvement over both in-vivo exposure therapy, which relies on access to real world stimulus, and in-situ exposure therapy, which relies on the patient's imagination (Koller, Schafer, Lochner & Meixner, 2019). In fact, this approach to exposure therapy has been popular enough to inspire the advent of a whole new area of treatment for issues such as phobias and PTSD, Virtual Reality Exposure Therapy (VRET). As the success of VRET is dependent on evoking realistic emotions, presence is considered to be highly important in this context and has been discussed by many researchers (Ling, Nefs, Morina, Heynderickx & Brinkman, 2014; Botella, Fernández-Álvarez, Guillén, García-Palacios & Baños, 2017; Price & Anderson, 2007; Koller, Schafer, Lochner & Meixner, 2019).

We see several commonalities in all of these experiments, whether discussing co-design or exposure therapy. First, they rely on realistic responses to the virtual environment in order to be effective. Researchers are interested in seeing how subjects would feel about or react to the contents of the virtual environment when faced with similar environments in the real world, making the transferability of behavior crucial. As we have seen, presence is the key element here, as it is what enables virtual environments to elicit realistic reactions, behaviors, and emotions. Second, the majority of these contexts involve asymmetric virtual reality setups. Whether one is the moderator of a usability study, a psychologist studying human behavior, or a psychiatrist administering VRET, they are playing the role of the external user. Meanwhile, their subject is the immersed user.

These two factors present a challenge for practitioners when combined. Maintaining the presence of their immersed users is essential, but due to the nature of their asymmetric setups, many of the interactions that they would normally rely on do the opposite. As we saw, researchers such as Koller, Schafer, Lochner & Meixner (2019), Zenner et al. (2018), and Liszio & Masuch (2016) have experimented with integrating interactions into virtual environments in order to circumvent this disruption in presence, with positive results. However, while Zenner et al. (2018)'s text-based communication and Koller, Schafer, Lochner & Meixner (2019)'s embodied avatars were effective in their own research contexts, the wide range of possible asymmetric virtual reality contexts seems to suggest that additional, alternative methods of integrating interactions are needed. Further, it is important for the designers of these virtual environments to understand how the method of integrating interactions between the two users affect the experience and dynamic between the users. In some situations, a more passive approach such as the text-based notifications may be of benefit, while in others, having the external user embodied in the environment might be more suitable. Our experiment aims to tease out these dynamics, as discussed in the following section.

2.6. Conclusion

As we have seen, maintaining presence during asymmetric virtual reality interactions is a crucial yet challenging aspect of allowing immersed users to interact with external users. We took a look at several attempts that have been previously made in an effort to avoid disrupting the immersed user's sense of presence by integrating interactions into the context of the virtual environment – for instance, by displaying notifications on a virtual television or having the moderator embody an avatar. We also explored contexts where such an asymmetric setup may be of use, drawing examples from design, psychology, and psychiatry. Going forward, we aim to expand on this current literature by exploring various methods of integrating interactions into the virtual environment, with the hypothesis that different methods of integration will have different effects on the immersed user and the dynamic that they have with the external user. Our goal here is to develop a nuanced understanding of what types the benefits and drawbacks of various interactions methods are, depending on the goals of the designer of a virtual environment.

3. Methodology

The previous sections identified two main research questions, based on the current literature on presence and interactions in asymmetric virtual reality, as being relevant to this research. Firstly, we aim to develop an understanding of *how different degrees of interaction integration affect the immersed user's sense of presence*. Secondly, building on that, we aim to understand *how the dynamic between the immersed and external user changes based on the degree of integration*. Using the literature and these research questions, we designed an experiment consisting of an asymmetric virtual reality environment suited for studying the effects of various types of interactions, each featuring a different degree of integration into the environment. This experiment was run with fifteen unique users, with each session consisting of a playthrough of the virtual environment followed by a post-session interview.

3.1. Interactions between the immersed and external users

The interactions we used in our experiment are categorizable under two categories: 1) verbal interactions, which allow the immersed user and external user to interact directly through speech, and 2) spatial interactions, which allow the external user to guide the immersed user through the virtual environment using environmental cues. Verbal interactions were chosen as one category due to their prominence in asymmetric virtual reality setups. Take two of our virtual reality usecases from the previous section, user experience research and psychiatry; both of these contexts often involve the moderator speaking to the user, guiding them and asking questions about their subjective experience. Spatial interactions were included due to their more subtle nature – an interesting contrast to the verbal help. While they are prominently featured in traditional video games, their effect on presence in asymmetric virtual environments is less explored, making them a worthwhile addition.

During the experiment, the immersed user completed several simple tasks with the help of the external user. This help was delivered through various interaction methods. Some of the interaction methods were designed to be highly disruptive to the immersed user's sense of presence, such as the external user speaking to them while next to them in real life, without any effort to integrate the voice into the virtual environment – a practice that is common in both entertainment and research contexts today. Others were designed to be more subtle, featuring higher degrees of integration, with the aim of preserving the immersed user's sense of presence. Examples of such interaction methods include speaking to the immersed user through a virtual object, such as a Walkie-Talkie, or pointing to objects using a spatial cue such as a beam of light. A full list of the interaction methods that were used during the experiment are outlined in Table 1. During the experience, we referred to these interaction methods as "conduits."

Table 1. List of conduits

Conduit	Description	Type	Level of integration
Video	The moderator appears on a virtual television inside the cabin. The user is able to see and hear them throughout the duration of the conduit.	Verbal and spatial	High
Walkie-Talkie	The moderator provides instructions through a virtual Walkie-Talkie. The user can pick up the Walkie-Talkie, press the button to speak to the moderator as well. The audio is adjusted to be spatial.	Verbal	High
Voice	The moderator provides instructions by speaking to the user through their virtual reality headset's headphones. The audio is not adjusted.	Verbal	Medium
Yelling	The moderator provides instructions to the user by speaking to them from next to them in real life, with no attempt to integrate the audio into the virtual environment.	Verbal	Low
Beam of light	The moderator highlights objects using a subtle beam of light shining through the window.	Spatial	High
Outlines	The moderator highlights objects using bright, glowing outlines.	Spatial	Low

List of conduits, or methods of interaction, that will be featured in our experiment. These allow the external user to guide and communicate with the immersed user.

3.2. The virtual environment and tasks

We chose to model our virtual environment, and the tasks that the immersed user completed inside of it, based on the concept of escape rooms. We hoped that most users would be roughly familiar with the concept of escape rooms, which have grown in prominence over recent years. This helped us communicate the session instructions to users succinctly and reduces the need for detailed instructions on what was going to occur in the virtual environment. By simply telling the user that they are in an escape room and need to complete five tasks to escape, we provided them with a straightforward mental model of what the session was going to be about without introducing a high level of cognitive load at the outset of the experiment.

Our virtual environment's theme was a simple log cabin. The 3D model of the cabin was modified to match the dimensions of our lab's open space in real life. This decision was rooted in the literature on locomotion in virtual environments, which suggested that movement mapped to real-life movement is preferable to flying or teleporting when it comes to user presence (Usoh, Catena, Arman & Slater, 2000). Similarly, design factors that were suggested to be conducive to maintaining presence, such as realistic shadows (Bowman & McMahan, 2007) and a highly interactable environment (Slater, Khanna, Mortensen & Yu, 2009) were integrated. Our goal in designing the virtual environment was to make sure that our chosen design factors were conducive to maintaining the immersed user's sense of presence – we didn't want the environment to be the reason for breaks in presence. By this same reasoning, we opted to go with a realistic environment rather than a science-fiction or fantasy theme. By the end of our design process, our log cabin was full of physics-based knickknacks and objects that the user could play around with, as well as environmental effects such as a crackling fireplace and snowy scenery outside.

Going along with our escape room theme, the tasks that the immersed users were solving were designed as simple puzzles – finding shapes around the cabin, figuring out the code for a lockbox, and so forth. The goal here was to design straightforward tasks that served as a backdrop for the interactions taking place. As such, the puzzles were specifically designed to require assistance from the external user (although they were possible to complete by one's self, if one knew where to look). A full list of puzzles that were included can be found in Table 2.

Table 2. List of puzzles.

Puzzle name	Description
Letters	The user must spell out a five-letter word ("GABRO"). The word can be found on various objects around the cabin.
Shapes	The user must find all four of the shapes. The shapes can be found around the cabin.
Snowball	The user starts out with a key. They must use it to open a window, take the snowball that they find outside, and use the stove in the cabin to melt it.
Blue numbers	The user must find all four blue numbers around the cabin.
Red numbers	The user must find all four red numbers around the cabin.

List of puzzles, or tasks, that the immersed user will be solving during the experiment. These provide an opportunity for them to interact with the external user in the context of semi-creative problem solving.

In order to distinguish between the effect of each interaction method, we designed the puzzles to be completed one at a time. This allowed us to assign a single interaction method to each puzzle before the session. Likewise, the order that the user completed the puzzles in, and the interaction methods assigned to puzzles, were randomized for each session.

3.3. Exploring presence

We explored the immersed user's sense of presence using both questionnaires and semi-structured post-session interviews.

Our questionnaire was based on questions selected from existing presence questionnaires. As seen in Schwind et al., we decided to integrate the questionnaire into the virtual environment itself, allowing us to prompt the user to fill it out after each task (Schwind, Knierim, Haas & Henze, 2019). This design decision was made in order to better understand how each task's associated interaction method affected the immersed user's sense of presence. Additionally, integrating it into the virtual environment allowed us to collect feedback without having the user exit the virtual environment between each task.

Our post-session interviews allowed us to collect rich qualitative data from users after the experience. During the interview, we asked probing questions about how they felt about each interaction method, whether they noticed changes in their sense of presence, whether anything in the environment struck them positively or negatively, and so forth. Users were encouraged to speak freely and to cover whatever came to mind at the time.

3.4. Procedure

Our experiment was expected to take about an hour and half in total. This includes the welcoming of the test subject, helping them put on the virtual reality equipment, the user session itself, and the post-session interview. The selection criteria for our test subjects included an age between 18-45 years old, a roughly equal gender balance, and that they do not professionally work in virtual reality. Participants were rewarded with two movie tickets (value of 20€) as an acknowledgment of their valuable contribution. The experiment was divided into four parts, outlined below.

3.4.1. Welcoming the participant and informed consent

Participants are guided to the lab upon arrival at the building. Once they have been shown into the space, they are provided with a consent form, which the moderators talk through with them. During this phase, users are informed of ethical and practical issues such as 1) the purpose of the test being to study the virtual environment and interactions that occur within it, and not their puzzle solving abilities, 2) that the user is free to ask questions or end the experiment at any point, 3) that we are collecting data from the sessions, including video recordings of both the physical space and virtual environment, as well as their questionnaire answers, and 4) that results will be anonymized and the identities of subjects will remain undisclosed.

Due to the ongoing COVID situation, special safety precautions were taken. Subjects are asked to wash and disinfect their hands upon arrival. The virtual reality gear is disinfected before and after every user session. The lab is left vacant for two hours between sessions, with windows left open to let the air circulate. Subjects are offered an optional facemask, and moderators wear facemasks for the duration of the session.

3.4.2. Introduction to the virtual reality gear

After the initial welcoming and consent, the user is introduced to the virtual reality gear. We let them try out the controllers ahead of time and explain where each button is located with the help of printed diagrams. Once ready, the user puts on the headset and picks up the controllers. The user is first presented with a tutorial virtual environment. Here, they can practice picking up objects, using number pads, and familiarizing themselves with the controller's various buttons inside the environment. Once the user is ready to proceed, they can initiate the experiment.

3.4.3. The experiment

The user is transported to the virtual environment, the log cabin escape room. After being given a moment to familiarize themselves with the surroundings, users are prompted to open one of the five lockboxes in the cabin. They are reminded that these lockboxes contain the puzzles. Once they open the first lockbox, the other four lockboxes lock until the current puzzle is completed. While solving the puzzles, the moderator engages the immersed user using the appropriate conduit and with help from them, the immersed user solves the puzzles. Upon completion of a puzzle, the user is rewarded with a key, which they can place in one of the five locks on the cabin's main door. This gameplay feature is intended to help the user conceptualize their progress throughout the experience.

After each puzzle, the television inside the cabin activates, showing our presence questionnaire. The user must complete the questionnaire before moving onto the next puzzle. This process is repeated five times. Once all five puzzles have been completed, the cabin's front door opens and the user "escapes" from the escape room. The screen fades to white, and the user is instructed to remove the headset.

During this time, one tester remains in the room with the user, while the other is in the back room operating the desktop build. Communication between the testers is handled via instant messaging on their smartphones to avoid disturbing the user.

3.4.4. Post-session interview and conclusion

After the session, the testers sit down with the subject and discuss the experience. After general questions about what they experienced, the testers prompt the subject to describe their reactions to the interaction methods. Careful attention is paid to how the subject describes the dynamic between themselves and the moderator. Once the interview concludes, the subjects are rewarded with the movie tickets and are escorted out of the building. The lab is disinfected and left vacant, as per safety guidelines.

3.5. Experiment setup

Our virtual environment is split into two builds, each featuring a different set of functionalities. The first build is intended for the immersed user and was be rendered inside the virtual reality headset. The second is for the external user and ran on a nearby desktop PC. The virtual reality build is designed for first-person interactions within the virtual environment, allowing the user to interact with objects and complete tasks. Meanwhile, the desktop build features moderator features similar to those outlined by Horst, Dörner & Peter (2018). These include the ability to initiate the session, controls for each interaction method, a bird's-eye view of the virtual environment for monitoring purposes, and emergency functions such as spawning lost items or directly communicating to the immersed user verbally in case they are stuck on a task.

The virtual environment was built in Unity, the popular cross-platform game engine by Unity Technologies. Virtual reality integration was handled through SteamVR, a software development kit released by Valve Corporation. Networking between the virtual reality build and desktop build was handled using Mirror, a high-level networking API for Unity. For video and voice streaming, we used the Unity plugin WebRTC Video Chat.

Our virtual reality setup consisted of the HTC Vive Pro 2.0 headset and two HTC Vive Controllers. We also used the Vive Wireless adaptor to enable untethered movement. Room-scale tracking was achieved using a pair of Vive Base Stations on opposite sides of the play-space.

3.6. Conclusion

Here, we have described the experiment and procedure that we have set up in order to explore the question of how various methods of integrating interactions into a virtual environment affect the immersed user's sense of presence, as well as the dynamic that exists between them and the external user. Our experiment consisted of five puzzles and five conduits, with the pairing randomized during each user session. Our goal was to see how various interaction methods affect the immersed user, and to that end, we were both recording their reactions during the sessions and conducting post-session interviews. In the next section, we explore the data generated from these sessions.

4. Findings

This section presents the findings of our user studies. These findings are primarily based on empirical data collected from the fifteen semi-structured post-session interviews, as well as comments made during the sessions. Emphasis was placed on data pertaining to how users felt about the conduits and the interactions that occurred during their usage. An anonymized list of participants can be found in the appendices.

During the analysis of user comments and interview answers, it became clear that rather than focusing on the degree to which a *conduit* is integrated into the environment, it is more insightful to focus on the degree to which the *external user* was integrated into, or embodied, in the virtual environment. That is to say, how much of a physical presence did the external user have during each conduit? For instance, during the spatial conduits the degree to which the external user was present was minimal – they were not embodied in the environment whatsoever, neither verbally nor visually. Meanwhile, during the Walkie-Talkie conduit, they had a physical anchor in the virtual environment, taking the form of a static object that the player could move around and interact with. While this difference may appear subtle, focusing on the degree of external user embodiment rather than the degree of conduit integration allows us to better align with how users conceptualized the interactions – the focus was almost always on the other user and how changes in conduits affected the dynamic of the interactions with them, rather than the conduits themselves.

Further, we found that the degree of conduit integration did not always lead to positive results. Take the case of the spatial conduits; there, attempts to increase the integration of spatial cues into the environment by focusing on visual realism often ended up detracting from the user's sense of presence, rather than increasing it as originally hypothesized (discussed later in this section). As such, simply focusing on the degree of interaction integration does not seem to be the most insightful approach when it comes to understanding what was happening during our user sessions. By focusing on the degree of external user embodiment instead, we are able to develop a more robust framework of how the interactions between the immersed user and the external user affected the dynamic between them, as well as the effect that they had on the immersed user's sense of presence.

When sorting the conduits by the degree of external user embodiment associated with them, we found a handful of consistent factors that went up or down depending on the degree of embodiment. Together, these factors help us understand the relationship between the immersed user and the external user, or moderator in our case, during the use of the respective conduits. These factors were: 1) the degree of control that the moderator has, 2) the degree of autonomy that the immersed user has, 3) the degree to which the immersed user is aware of the moderator, and 4), the degree of social feedback that the immersed user receives from the moderator. As we will see, variations in these factors have a significant effect on how the immersed user perceives the moderator's role in the virtual environment, and in turn, whether the interactions with the moderator affected the immersed user's sense of social presence or spatial presence. This relationship between external user embodiment, the factors affected by the degree of embodiment, the resulting perceived role of the moderator, and the components of presence being affected by the interaction are visualized in Figure 1.

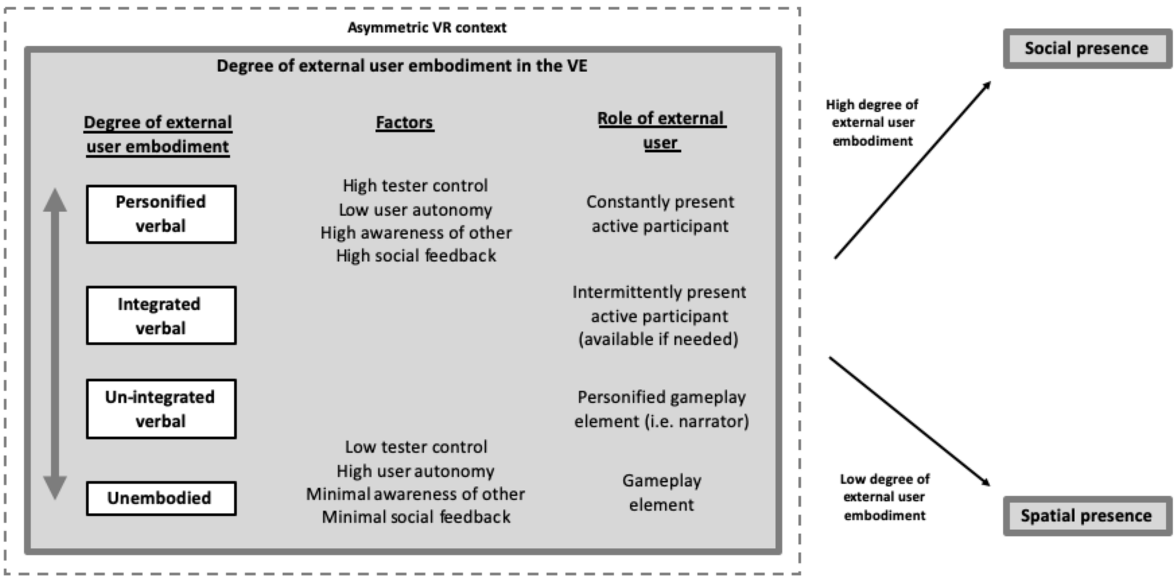

Figure 1. Relationship between degrees of external user embodiment, influenced factors, and the perceived role of the external user.

The following sections will further elaborate on this framework and the components that make it up, beginning with the degree of external user embodiment, followed by the factors influenced by the degree of embodiment and their effect on the perceived role of the external user, and finally, concluding with how the immersed user's sense of presence changed depending on the degree of embodiment and associated factors.

4.1. Degree of external user embodiment

The degree of external user embodiment describes how integrated into the virtual environment the external user is. In our experiment, this integration was achieved visually, in the form of virtual objects representing the external, and auditorily, by including the external user's voice in the virtual environment. The degree of embodiment comes on a spectrum, ranging from no embodiment in the virtual environment at all to being highly embodied.

Our two spatial conduits provide an example of a minimal degree of external user embodiment. During these conduits, the external user had no discernable presence in the virtual environment. While they were able to observe and interact with the immersed user using the Desktop build, the immersed user had no indication that this external user participation was occurring. This often resulted in our test subjects assuming that the external user's actions in the virtual environment were simply gameplay mechanics, rather than the actions of another user. This minimal degree of embodiment, where the external user is participating through the environment rather than actually being represented in the virtual environment, can be referred to as *Unembodied*.

Going up from there, you have increasing levels of external user embodiment, each defined by a significant shift in how clearly the external user is represented in the environment. One degree above the external user only participating through spatial cues, you have *Un-integrated verbal embodiment*. Here, the external user is represented through verbal communication. While the external user is not anchored to the environment through any visual means, the immersed user is able to hear their voice, providing a significantly higher degree of personification than seen with the spatial conduits. The next conceptual shift in embodiment occurred when users were provided with a physical representation of the external user. Now, not only could the immersed user hear them, but they were physically integrated into the virtual environment – leading to the classification of *Integrated verbal embodiment*. Finally, you have the highest degree of external user embodiment, *Personified verbal embodiment*. Here, the external user themselves is represented in the virtual environment – not just as a static object, but as a fully embodied character or avatar. This degree of embodiment is differentiated from the previous one based on the external user being fully personified in the environment - the immersed user is not interacting with some remote person through a virtual device, but rather, is sharing the space with a representation of a human and is experiencing all of the social factors that come with that (as discussed in section 4.2.4). A full list of conceptual degrees of embodiment can be found in Table 3.

Table 3. Conceptual degrees of external user embodiment.

Degree of external user embodiment	Description	Example conduit designs
Personified verbal embodiment	The external user is embodied as an avatar or character, sharing the virtual environment with the user. Interactions are both verbal and spatial.	*Video* conduit; Koller, Schafer, Lochner & Meixner (2019)'s embodied audience members
Integrated verbal embodiment	The external user is embodied as a static shape such as a loudspeaker or phone. Verbal interactions are remote and intermittent.	*Walkie-Talkie* conduit
Un-integrated verbal embodiment	The external user is not embodied in the virtual environment, but the immersed user can hear and speak to them verbally.	*Voice* conduit
Unembodied	The external user is not embodied in the virtual environment at all. Interactions occur through spatial conduits rather than voice.	*Outlines* conduit; *Beam of Light* conduit; Zenner et al. (2018)'s video screens

4.2. Factors

Based on our data, we identified several key factors that were determined by the degree of external user embodiment. These factors influenced how the immersed user felt about the role of the external user and the dynamic that existed between the two. They are explored below.

4.2.1. Degree of moderator control

The first factor that was significantly affected by the degree of external user embodiment was the degree of control that the moderator had during the session. We found that the less embodied the external user was, the less control they had over the actions of the immersed user. For instance, it quickly became clear that the spatial conduits offered the moderator a significantly lower level of control compared to the various verbal conduits. While the moderator was able to activate and deactivate the conduits upon their own discretion, it was up to the user to notice and correctly interpret the assistance. Often, this was a smooth, nearly instantaneous process - the user would look around the virtual environment, see an outlined object, and immediately understand it to be the puzzle piece that they are missing. However, occasional but consistent issues arose in either the noticing or understanding phases. Noticing a spatial cue depended on the user looking in the right direction and actively searching for things that would help them on the current puzzle. During these situations, the moderator would be aware that the user had to turn around to see the active spatial cue but was unable to instruct them to do so. Similarly, the user may have spotted and gazed at the active spatial conduit, but interpreting the meaning was still up to them. This moderator passivity had an effect on the user, as discussed in the next section, user autonomy.

Compared to the spatial conduits, our verbal conduits offered the moderator an immediate way to offer the user assistance when needed. This was especially the case with the Voice and Video conduits, where the moderator's decision to speak to the user was entirely up to them; when the moderator spoke to the immersed user using these two conduits, the user had no choice but to receive the verbal cues. Interestingly, this was not the case with the third verbal conduit, the Walkie-Talkie. Here, we observed the conduit providing the immersed user with an unforeseen level of control over the interaction when compared to other verbal conduits. With the Voice and Video conduits, it was up to the moderator to decide when to engage the user. However, with the Walkie-Talkie conduit, we were surprised to find that some users decided to not respond to the incoming call – and because of the nature of the conduit, this resulted in the moderator being unable to interact with the user. While not common, this occasionally occurred when the user in question was engaged by the puzzle they were solving and wanted a chance to figure it out by themselves before accepting assistance. One user expressed this sentiment very clearly as they neglected the beeping Walkie-Talkie, exclaiming "*Yeah, yeah, I know, I know, but I'm not gonna answer. I can do it by myself.*" [Participant 9] Another user commented on this during the post-session interview, stating that they had been "*very tempted to ignore the call*" before adding: "*I was really resisting the temptation of calling you, but in the end I was like ok, we need to get this thing over with.*" [Participant 6]

Compared to the spatial conduits, the verbal conduits not only provided the moderator with an immediate way to reach the immersed user, but also enabled a far richer and more versatile medium for interaction. Often, this turned out to be necessary when answering user questions or providing instructions on how to recover from an unexpected situation (such as a glitch). Indeed, when commenting on the verbal conduits in general, several users expressed feeling a sense of comfort knowing that the tester could be reliably reached and relied upon to provide help if needed.

4.2.2. Degree of user autonomy

The second factor, the degree of user autonomy, is closely related to the degree of moderator control. In fact, the two seem to have an inverse correlation – when the moderator has a high degree of control over interactions, the immersed user is left with a reduced level of autonomy, and vice-versa. This is clearly illustrated with the spatial conduits, where the moderator had a minimal amount of control over whether the user would notice and understand the visual cues in the environment; when not receiving verbal guidance, the user was left to explore the space at their own pace, allowing them to better take in the visual and auditory elements around them. While they solved the puzzles less efficiently, they were reportedly more connected to the environment around them. This is illustrated by one user's succinct comment on the topic: *"yeah I was alone, so I could hear the bird which I didn't hear before. And the fire. And at some point I was feeling like I should sit down at the table. And the window! The window scene was…. I know it was VR, but I almost felt like uuuu it's cold."* [Participant 8]

Meanwhile, when the moderator had a high degree of control and was able to interject with verbal guidance at will, the immersed user was left with a reduced level of autonomy. They were no longer exploring the virtual environment by themselves, but rather, being guided through it.

4.2.3. Degree of awareness of the moderator

This factor deals with the question of how aware of the moderator's presence the immersed user was. As was the case with the others, this factor was highly dependent on the degree of external user embodiment – the more embodied the external user was in the virtual environment, the more consistently aware of their presence the immersed user was.

This is best illustrated by comparing the two extremes, the Video conduit and the two spatial conduits. In the case of the former, the moderator appeared on a large screen in the virtual environment. The video feed was left on for the duration of the conduit interaction. The immersed user not only had a clear visual representation of the external user's presence in the environment with them, but they also had a visual indicator that their actions were constantly being watched by said external user. While users were being closely monitored throughout the entirety of each session, it was only during this conduit that users expressed feeling that way. As one user put it, *"it felt like you were watching. But then again I guess you were watching the whole time anyways. But yeah that definitely felt like, real, like you were watching everything I was doing. And that always feels a bit weird."*. [Participant 15] Overall, users reported feeling hyper-aware of the external user's presence, illustrated by comments such as:

"But when you showed up on the screen with me that was a bit different because then I had… you know… I was very aware of your presence at that time."; [Participant 9]

"Like it felt like, it felt like you were watching me with that TV." [Participant 3]

"Yeah I didn't like that one. I think it's… because you're in that space, when you'd rather just… hear a voice. It's kind of weird, oh don't come and see me, I'm in this.. this cabin now. Don't intrude on my experience!" [Participant 10]

"Yeah seeing your face actually…. I feel like instead of creating closeness it felt like I was under some onus to take help, so I'll reduce the rating." [Participant 9]

The effect that this high level of awareness had on the immersed user is discussed in the next section, which covers the perceived role of the external user and the dynamic between the two. Meanwhile, if the Video conduit encouraged hyper-awareness of the other user, the Spatial conduits did the opposite. While the user may have asked a few questions that received no responses when first starting with these conduits, they quickly adjusted to solving the puzzles by themselves with a high degree of autonomy. When the users did receive help in the form of spatial cues, they often didn't associate the assistance with the moderators at all – instead, they thought that the Outlines or Beam of Light were simply gameplay elements or environmental effects, not cues provided by another person. Many users expressed this same sentiment:

"*I thought that it was just making the environment more interactive, having sunlight coming through. I didn't think that it was a hint from you*"; [Participant 4]

"*Ohhh that's true, I thought that it was part of the…. design of the game. Like you play a game and certain objects blink or glow, and that's very different to talking to someone about how to solve this level*." [Participant 9]

This too is further explored in the following discussion on how these factors affected the external user's perception of the role of the external user.

4.2.4. Degree of social feedback

This last factor deals with the degree to which the immersed user received feedback to the fact that they're being listened to when speaking. While the users were obviously always heard by the moderators, whether or not the users felt that way depended on how much the active conduit allowed for social feedback from the external user. Take for instance the Video conduit; the user had a clear visual indicator that they are being monitored and listened to. Meanwhile, with the Spatial conduits, no such indicator existed – and indeed, users reported not being sure whether anybody was listening to them as they spoke aloud during those interactions: *"Because I remember that I was like in my mind, constantly thinking, it's so awkward to just scream or shout that I need help, cause I made some comments but there was no feedback, so I was wondering, are you guys listening or not?"* [Participant 3] While this uncertainty was reduced with the Voice conduit, as most users seemed to simply assume that they were being heard, several users still commented on the lack of feedback here too: *"Maybe that's also why I didn't know how to ask for help. But I think I was more comfortable to converse when I knew that there was someone actually listening to me…. I don't know if in the call, there is some way of just having some sound or noise or something to show that the other person is on the line"*. [Participant 4]

The Walkie-Talkie conduit provided an interesting middle ground when it came to social feedback. During post-session interviews, users reported enjoying the physical interaction associated with using the Walkie-Talkie: going over to it, picking it up, pressing the button on the side, and holding the button to speak. While the Walkie-Talkie conduit provided less visual feedback than the Video conduit, this series of actions offered the user an affordance and feedback mechanism. The user knew that as long as the Walkie-Talkie was active (indicated by the green LED light on top), if they held down the button and spoke, the external user would be able to hear them and would respond after they let go of the button.

4.3. The perceived role of the external user

The previously discussed factors, determined by the degree of external user embodiment, appeared to have a significant impact on how the immersed user perceived the external user and their role in the virtual environment. Depending on the degree of embodiment, the user's perception of the external user's role ranged from not perceiving the external user at all to feeling that they are an active collaborator sharing the virtual space with them. This perception of the external user's role significantly affected how the immersed user responded to and interacted with the external user.

Based on the data collected from the user sessions and post-session interviews, there were four primary ways of conceptualizing the role of the external user: 1) as a gameplay element, 2) as a personified gameplay element, 3) as an intermittently present active participant, and 4) as a constantly present active participant. These conceptualizations are discussed below.

4.3.1. The external user as a Gameplay Element

During the spatial conduits, the external user saw the lowest degree of embodiment in the virtual environment – represented neither visually nor verbally, unable to interact with the user beyond the use of the two conduits to highlight objects. This resulted in a high degree of autonomy for the immersed user, along with a minimal amount of social feedback and awareness of the other. As a result of these factors, the immersed user often did not perceive the external user as being there at all. In fact, as mentioned when discussing the awareness factor, even when the immersed user received help from the moderator, it was common for them to not perceive it as assistance from another person. Instead, user often conceptualized the help that they were receiving as being gameplay elements – simply another pre-programmed part of the escape room that they were playing through, like the puzzles, lockboxes, and keys. Several examples of responses were included when discussing the user awareness factor; a few others include:

"I didn't pick up on the fact that them glowing was you trying to help me." [Participant 9]

"Maybe the help was there but I didn't realize it was a cue from you guys, I thought that it was just me being smart…" [Participant 13]

"I mean one thing I noticed at the end, I'm not sure if that was… I guess that was something that you guys did as well, when I was looking for the last wooden box to put into the frame, it was flashing a bit. Not sure if that's just me approaching the item or actually you guys…. you know, lighting it up." [Participant 6]

"I do remember it was flashing, but I didn't think of that as being a cue from you guys, I thought it was just that when you look at it, it automatically flashed or something like that, like it was a reflection or something." [Participant 3]

"Now that I think about it, in hindsight, I can see that it was… it wasn't a real thing, it was a clue." [Participant 11]

"Though finding where the key goes was interesting until, I don't know if you activated it or if it's an automatic thing, but the window started glowing, so I was like ok, that makes sense" [Participant 1]

"I thought that it was blinking because maybe I'm running short on time, I thought in this way." [Participant 4]

4.3.2. The external user as a Personified Gameplay Element

During the Voice conduit, we saw the external user's embodiment limited to voice, with no visual or physical presence in the virtual environment. While the moderator had more control in this situation, the social feedback that the immersed user received was limited and their awareness of the external user was not reported as being high. Indeed, when asked about their conceptualization of the external user during this conduit, users described them using terms such as "narrator" or "guide." One user likened the interaction to receiving military-esque commands: *"I thought that it was just like in a military environment, you just receive commands from the headset, you don't talk back to the headset, but in this case with the video call, I think I was comfortable talking back and asking questions."* [Participant 2] Another described the voice they heard as part of the gameplay: *"And of course, just* *speaking to me through the headphones. Uh… almost feels like a video game, like… a narrator or something, talking straight into your ears, into your mind."* [Participant 15]

This sentiment was echoed by the behavior of users – unlike with the two other verbal conduits, users often did not directly respond to the external user at all when using the Voice conduit. This is a significant deviation from their behavior during the Video and Walkie-Talkie conduits, where users often replied with a "thank you" or some other form of acknowledgement. This suggests that the immersed users did not consider the external user to be an actual participant, adopting a more transactional, one-way relationship. At the same time, unlike with the spatial conduits, it is clear that the immersed users understood the Voice conduit to be associated with a person, rather than simply being an environmental effect. Considering this, the description of the external user as a narrator seems apt.

4.3.3. The external user as an Intermittently present active participant

Once a physical representation of the external user was introduced into the virtual environment, we saw a shift in the immersed user's behavior towards them. During both the Walkie-Talkie and Video conduits, the immersed user went from seeing the external user as part of the gameplay to being an active participant in the virtual environment, to "being there."

In the case of the Walkie-Talkie, the external user was there if needed – as discussed in the section on moderator control, the user not only controlled when the interaction began on their side, but also had the option of not accepting the call when the external user was trying to get in touch. This created a sporadic relationship, rather than a constant one as seen with higher degrees of embodiment.

During this conduit, the immersed users described the role of the external user in a positive light, often focusing on the cooperative nature of their dynamic: "*Yeah the walkie talkie is the one where you, there's this one video game where you're put in separate places and you talk to the other person with the walkie talkie. And for me that was the one that I felt, ok this person wants to help me. And not just watching somewhere."* [Participant 10] Users commonly spoke about the external user being physically present in the virtual environment, despite their presence being limited to a static object. This conceptualization is illustrated by several examples:

"I'd say the walkie talkie one was probably the most immersive. […] But the walkie talkie one, you know, it's like you're somewhat in the virtual environment talking to me. That's how it felt like. [Participant 15]

"*It was more like two-way. For me it was the most fun part. I had to do something, and you were there. The visual, it felt a bit different from the previous sound just coming from my headset.*" [Participant 8]

4.3.4. The external user as a Constantly present active participant

Similar to the previous category, here too the immersed user considered the external user to be an active participant. Some described the external user as "being there", despite the external user's embodiment being limited to a video feed on the virtual television. This dynamic was captured by several users:

"*I think that it brought me closer into the space that I was operating in, that you were there as a physical version of… of course not physical, but you get what I mean, like there was an artifact inside there with you. In that way it felt more real."* [Participant 5]

"No, he wasn't exactly with me in the cabin, but he was on my screen in the cabin, so he was there." [Participant 9]

Regardless of the external user's physical location, users unanimously agreed that the combination of verbal and visual social feedback created a strong sense of somebody else being there with them, and often, a sense of connection:

"But I liked the video one as well. That was quite interesting, cause it brought a connection – that was the most immersive part. You actually felt connected to whoever was there with

you. Like it felt like, it felt like you were watching me with that TV. And I could talk to you, and you could make eye contact, even though we weren't in the same space physically. So I think that that one was the most immersive." [Participant 3]

"It didn't feel as collaborative as the video felt for example, because in that one you were constantly there and it felt like, like a partner would have done in that situation, actually would have been in the room and helped you out." [Participant 10]

"I don't get scared really easily, but I could imagine that for someone that actually could get frightened in a situation like that – cause it is a little bit like intimidating, to be in like virtual reality, so I think that if somebody was actually scared then that [video] *could help them feel more comfortable in it."* [Participant 7]

"You get two sensory inputs, you get the visual, and you get the audio also. Obviously like that helps a lot." [Participant 6]

However, as mentioned in our discussion on the immersed user's awareness of the external user, this level of embodiment resulted in many users finding themselves hyper-aware of the other person. This led to a significant portion of users feeling uneasy about the interaction, often describing it as being watched or monitored, as seen previously. Naturally, this led to quite a different conceptualization of the role of the user than we saw with the Walkie-Talkie conduit, which was commonly positive. Here, users likened the external user to a boss and captor:

"Yeah, because at that point your role wasn't the helper who like helps you out when needed, it was actually like the captor in my mind." [Participant 10]

"*Someone is seeing you, that was one of the factors that influenced it, like ok, he's seeing, but with the audio help I felt like he's somewhere away, he's not here in this physical room; but when I see you on the screen, it's like you're there, you're there like a boss. Boss is looking at me…"* [Participant 5]

In one less dramatic conceptualization, the external user was compared to a game counselor: *"Will do. You seem like a… Nintendo game counselor right now. "* [Participant 9]

In these cases, the language being used is indicative of authority. Of course, it is important to note that we opted to use a video feed of the external user displayed on a television, rather than a virtual avatar, for this fully embodied conduit design. It is very possible that user sentiments would have been different had the design been different.

Regardless of positive or negative sentiment towards the external user in this case, it is clear that a big differentiator between this conceptualization and the previous one is the duration of and control over the interactions between the users. In the case of the former, the interaction was intermittent – the external user was not necessarily always there, but rather, was accessible when needed. In the case of this conceptualization, the immersed user is receiving constant, strong visual feedback to the fact that the external user is there constantly throughout the duration of the conduit. This changes the role of the external user quite significantly, going from somebody who can be reached out to for assistance to a very present, constant, second participant.

4.4. Effect on presence

Originally, our first research question was asking about the effect that integrating interactions into the virtual environment has on the immersed user's sense of presence. As we saw, it was necessary to refine the initial part of this question, reconceptualizing interaction integration as external user embodiment in order to develop a more nuanced understanding of the dynamics at hand. Similarly, we must now refine the latter part – it turned out that the degree of user embodiment did not simply increase the amount of presence that the immersed user experienced, but rather, changed the component of presence that was being acted upon. As seen in the literature, presence can be broken into three sub-components: social presence, spatial presence, and self-presence.

Depending on the level of external user embodiment, which subsequently affected the immersed user's perception of the external user's role in the virtual environment, the component of presence being acted upon varied between social presence and self-presence. During conduits with a higher degree of external user embodiment, the component being acted upon was clearly social presence. This is clear, looking at both the language that users used when describing the external user (as seen in the previous section) and the factors at play: high degrees of social feedback, high awareness of the external user. Likewise, conduits with a low degree of external user embodiment led to an emphasis on spatial presence. As we saw, when the spatial conduits were active, users were more engaged with the sounds and visuals, the ambiance, of the virtual environment. They spent more time actively looking for clues rather than following verbal hints. The third component of presence, self-presence, did not seem to be affected by the choice of conduit. This was to be expected – as suggested by the literature, self-presence is more dependent on non-asymmetric factors such as the immersed user's avatar and the quality of tracking (discussed in the following section on mediating factors), which remained standard across conduits.

While the component of presence being positively acted upon depended on the degree of external user embodiment, reductions and breaks in the immersed user's sense of presence were seemingly caused by similar causes, regardless of whether the immersed user was experiencing social or spatial presence. These causes were: 1) a lack of realism and 2) something drawing the user's attention back to the real world. The first cause, lack of realism, was seen when users noticed something about the virtual environment that seemed "off" or unrealistic. This aligns with Slater & Steed (2000)'s findings on reasons for breaks in presence. The second, the user's attention being drawn back to the real world, occurred when the user became aware of stimuli that was not mediated through the virtual environment.

Two conduits that commonly featured such breaks in presence were Yelling (speaking to the immersed user from the real world without any mediation) and Beam of Light (using a beam of light to point at objects within the virtual environment). For the former, it was common for users to report that speaking to them from the real world took their attention away from the virtual environment. Examples of user comments here include:

“So if I were to rate those and put those on a scale then… you talking right next to me… and in terms of establishing that sort of connection… you being in the same room but not interacting with me through the environment itself, but actually just talking to me and not even through the headphones, to me that was kinda… the least believable.” [Participant 6]

“You talking to me right next to me, that didn’t really uh… I didn’t want to have that in the experience; if I could have chosen, I would rather be without that.” [Participant 1]

“I guess that in a way it was a reminder that the real world, that you’re talking to me from the real world, whereas when it was through the headphones it felt like you were in that world with me.” [Participant 13]

“When I’m just hearing someone’s tips or words without having that kind of an interactive medium in the VR, it doesn’t feel that great, or maybe it’s even annoying, that I just hear from somewhere that ‘hey you should do this’, but it doesn’t feel like it’s within the… virtual world itself.” [Participant 12]

"For a second I was like what, but then it’s like yeah of course I’m still here." [Participant 9]

As the Yelling conduit was intentionally designed to be disruptive, this was not surprising. In the case of the latter, however, the breaks in presence were unintended. We had originally hypothesized that integrating spatial cues into the environment would be less distracting than using traditional video game techniques (as seen in our other spatial conduit, Outlines). Despite using a professional-grade Unity asset for the light effect, users did not seem to find the level of realism convincing. In fact, they found it more distracting than Outlines, which made no effort to be realistic. User comments here included:

“It was fun, but it was like oh, oh yeah, of course…. But it immediately reminded me that it’s a game, a virtual environment.” [Participant 14]

"Hahaha, I think that the god light was funny, but annoying at the same time." [Participant 10]

"And then the, like the, light coming from window, it was also… I think it was also a bit too bright, so then it wasn't that realistic." [Participant 2]

4.5. Questionnaire answers

Our virtual environment featured a short questionnaire that we asked the subject to fill out after each puzzle-conduit pair. The purpose of this questionnaire was to catch data on how the subject was feeling immediately after they had interacted with the external user through one of our interaction methods. We primarily used it to guide post-session interview questions and conversations, rather than as a direct method of analyzing user presence.

The reasons for this were twofold: first, while the questions were drawn from previous literature, the questionnaire itself was kept short in order to avoid five prolonged interruptions throughout the user session. As such, repeating questions to validate user answers were not included. Second, the behavior that users exhibited while answering the questionnaire was questionable. Often, they were in a hurry to get back to solving puzzles, resulting in what appeared to be rushed answers. It was also relatively common for users to contradict scores they had given when later discussing the issue during the post-session interview. Likewise, users would provide spoken reasons for their scores while answering the questions, such as their headset becoming loose during the puzzle or them not noticing any help or interaction (despite giving a score of 3 out of 5 for how strong of a connection they felt with the moderator during said puzzle). This nuance was not captured in the questionnaire scores. As such, the scores themselves were not taken into direct consideration when exploring our user data, instead simply acting as prompts for discussion in the post-session interviews.

4.6. Mediating factors

Thus far, this section has been focused on findings related to how the immersed user's sense of presence was affected by asymmetric interactions. While this will remain the primary focus for the remainder of the study, it is important to point out that there are many factors not related to asymmetric dynamics that influenced our user's sense of presence during our experiments. In fact, due to the open-ended nature of our post-session interviews, the bulk of comments that were made and analyzed had to do with these mediating factors. The effects of these mediating factors predominately aligned with previous literature on presence. Several of the most commonly brought up mediating factors are outlined below.

4.6.1. Physics interactions

Throughout our virtual environment, we had many physics-enabled objects that users could pick up and play around with. We found that when objects behaved how users would have expected them to, reactions were positive – the users expressed positive sentiments in response to physics that worked well. For instance, it was common for users to ask whether they could pick something up, try, see that they could indeed do so, and then say "*cool*!" or "*neat*!" This is in-line with Schuemie, Straaten, Krijn & van der Mast (2001)'s findings on the positive results of providing users with a high degree of ability to participate in and modify the virtual environment around them. At the same time, when objects did *not* react how users expected them to, users expressed negative sentiments. For instance, several users attempted to place items on the stove in the cabin, expecting them to catch on fire. In the post-session interviews, they mentioned that these instances reminded them that they were in virtual reality.

Our choice of locomotion, mapping user's movement in the virtual environment to their movement in real life, was also met with positive responses. Users made comments such as "*So I can actually move around this space? That is very cool.*" and "*I have to say that in terms of walking and moving distances and grabbing stuff, it does feel really correct in the sense that the amount of steps I'm taking, the distance I feel like I'm going in reality does* *feel realistic in this environment as well."* [Participant 12] This too aligns with Usoh, Catena, Arman & Slater (2000)'s suggestions on locomotion methods.

4.6.2. Visuals and audio

Across the board, users expressed positive sentiments towards the visual and audio elements of the virtual environment. One user succinctly summarized the commonly mentioned factors here: *"I appreciate the fireplace and how it was making the crackling sound, I appreciate the snow when you looked outside and you could see scenery, the lighting, the vibe, it all created this really cozy cabin feeling*." [Participant 6] As discussed earlier, these comments on the environment were most common when the user was left alone during the spatial conduits, allowing them to take in the ambience. One user captured this phenomenon, stating: "*When I was alone? Yeah, it was quite nice when everything was so quiet. I heard the fire, the birds, you know, real cottage experience. It was quite nice*." [Participant 8]

4.6.3. Virtual reality equipment

Occasionally, we had moments where the virtual reality equipment resulted in breaks in presence. These moments were aligned with Riches, Elghany, Garety, Rus-Calafell & Valmaggia (2019)'s suggestion that awareness of the virtual reality equipment can lead to breaks in presence. Several times, users were caught in the cable connecting the headset to the battery pack in their pocket. In other instances, users had to pause in order to tighten the headset. One user commented on this during the session, stating: *"so my presence actually decreased because my set is starting to become a bit loose… So, I need to tighten it. My vision is getting blurry."* In the post-session interview, they added: *"I thought it was valuable knowledge that the headset was becoming loose, so you felt less immersed, but you obviously can't know that the headset is loose when you're looking at a data set."* [Participant 9]

More frequently, we had moments where the virtual reality controllers caused the user to shift their attention from the task at hand to trying to figure out which button they had to press. One user described the effect that this hand on their sense of presence: *"remembering* *which button to press to perform which action… you do kinda get a hang of it, but especially in the beginning, when you're just figuring things out, that's kinda the… you know, most difficult in terms of immersion."* [Participant 6] Another pointed out that it reminded them that they were not in real life: *"I was having trouble pressing the buttons; for me that was a little frustrating, it was like, obviously if it were a real box, I could do this."* [Participant 7] Fortunately, several users reported that the initial tutorial room helped circumvent these issues somewhat, providing them with a chance to learn how to use the controllers before the actual experiment. As one user put it, *"the tutorial was really helpful. If you would have straightly put me in that room, it would have taken at least ten minutes for me to figure out what am I doing."* [Participant 5]

5. Discussion

This thesis set out to explore the relationship between the degree to which interactions are integrated into a virtual environment and the sense of presence felt by the immersed user in asymmetric VR environments, as well as how the degree of integration affects the dynamic between the immersed and non-immersed users.

5.1. Key findings

Based on our data, we found that the degree to which the external user was embodied in the virtual environment changed how the immersed user perceived their role in said virtual environment. This in turn determined whether the interactions between the users affected the immersed user's social presence or spatial presence.

We explored several factors that acted as intermediaries between the degree of external user embodiment and how immersed users perceived the role of the external user. These factors were: 1) the degree of control that the external user has, 2) the amount of autonomy that the immersed user has, 3) the degree to which the immersed user is aware of the external user, and 4) the degree of social feedback that the immersed user is receiving. As we saw, the more embodied the external user is, the more prominent their perceived role in the virtual environment is; when highly embodied, the immersed user is more aware of the external user's presence and receives more feedback to the fact that they are being listened to, while at the same time being less focused on exploring and absorbing the environment. Likewise, the less embodied the external user is, the more focused on the environment the immersed user is, but at the same time, the less aware of the external user they are.

We also saw that the degree of embodiment affected how much control the external user had over the interactions. With a high level of embodiment, they were able to directly instruct the immersed user and had much more control over the interactions between the two. With lower levels, they had less control, instead relying on subtle nudges to guide the immersed user, with the success of the interactions ultimately depending on the immersed user. We see examples of this playing out in the studies conducted by Koller, Schafer, Lochner &

Meixner (2019) and Zenner et al. (2018). When using Zenner et al. (2018)'s unembodied interaction method, notifications on virtual television screens, the external user has little control over whether the immersed user sees the messages. Meanwhile, Koller, Schafer, Lochner & Meixner (2019)'s highly embodied audience members provided the external user with the ability to directly instruct and interact with the immersed user.

5.2. Implications

Previous studies provided examples illustrating that by integrating interactions into the virtual environment in an asymmetric virtual reality context, one could avoid negatively affecting the immersed user's sense of presence (in fact, such interactions were shown to be conducive to high levels of presence). As discussed, Zenner et al. (2018) demonstrated this with integrated notifications on television screens, and Koller, Schafer, Lochner & Meixner (2019) with embodied audience members during a public speaking exercise.

Here, we moved past single examples of integrating interactions, instead examining how different types of integrations affect the immersed user's sense of presence and the dynamic between the two users. We showed that integration can be thought of as existing on a spectrum, with different degrees of integration resulting in different outcomes. Further, we have also shown that the type of integration affects the component of presence being acted upon rather than simply increasing presence as a whole, providing a more nuanced approach to understanding how the immersed user's sense of presence is being affected by asymmetric interactions.

This thesis adds to the larger body of virtual reality presence literature by demonstrating how external user embodiment affects an immersed user's sense of presence in an asymmetric virtual reality context. Previous literature has established presence as being made up of three primary components: social presence, spatial presence, and self-presence. Numerous authors have explored how different factors, ranging from virtual physics to hardware, influence these components in a standard virtual reality context. As we have seen, introducing an asymmetric element to virtual reality setups introduces additional factors that influence the user's sense of presence. This thesis explores how the degree of external user embodiment affects these additional factors, and how they in turn influence the social and spatial components of presence. The positioning of this thesis in the larger body of virtual reality presence literature is illustrated in the diagram below, Figure 2.

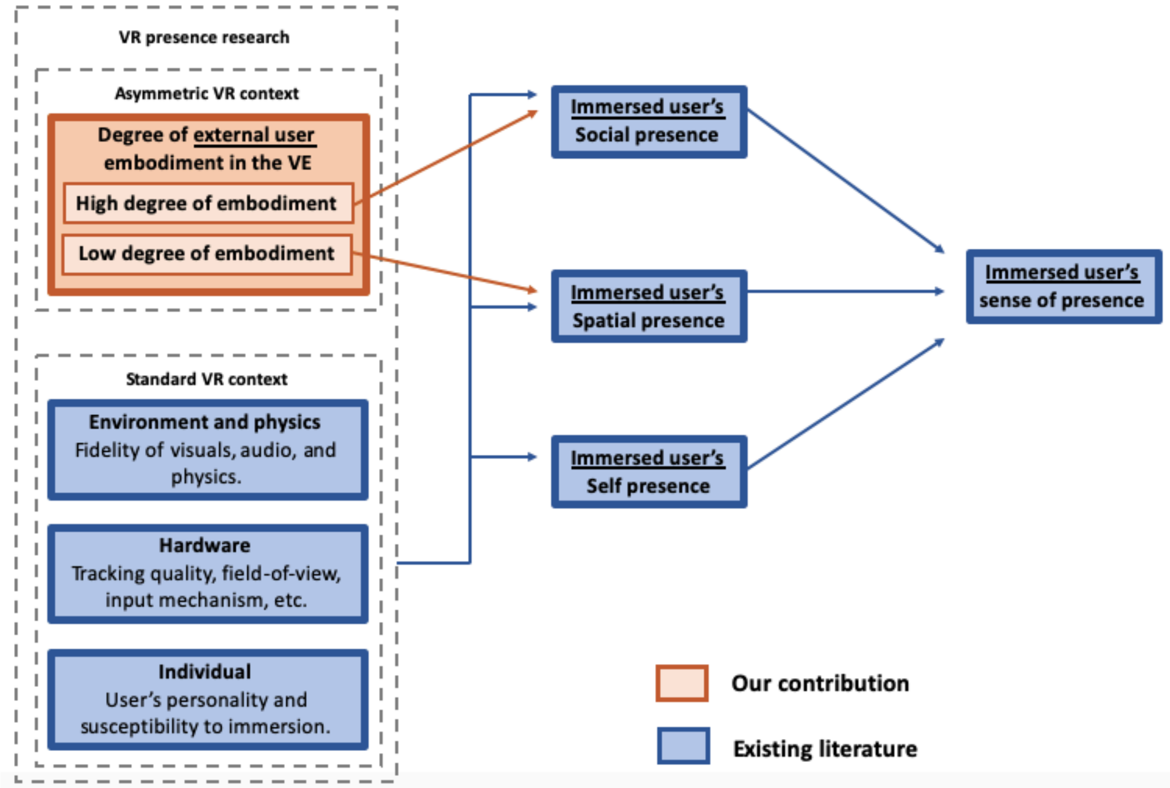

Figure 2. Positioning of thesis findings

6. Conclusion

This thesis set out to explore how different levels of integrating interactions into virtual environments affect the immersed user's sense of presence in asymmetric virtual reality setups, as well as how the dynamic between the two users changes based on said degree of integration. These questions are becoming increasingly relevant as virtual reality sees increasing adoption in numerous non-entertainment contexts, often featuring asymmetric setups with one user acting as the moderator administering the experience while the other is immersed in the virtual environment.

To this end, we set out by establishing our two primary research questions:

RQ1: In asymmetric virtual reality, how do varying levels of integrating interactions into a virtual environment affect the immersed user's sense of presence?

RQ2: How does the dynamic between the immersed and external user change based on the degree of interaction integration in asymmetric virtual reality?

These questions were inspired by the current state-of-art literature on the topic of presence in asymmetric virtual reality. As explored in section 2, asymmetric virtual reality is a context that is seeing increasing use in a variety of fields.

One primary example of this was in the context of co-design and user experience research, with practitioners using virtual reality as both a method of immersing subjects in a different context and for cost-effective prototyping. Several examples of this were presented in section 2.5, ranging from the design of hospitals (Tiainen & Jouppila, 2019) to pharmacies (Mobach, 2008) On the context-immersing side, this approach has several advantages over traditional methods such as design artifacts or narratives, including availability and practicality, cost, and iteration time. Similar benefits are seen when it comes to prototyping; by adopting virtual reality, designers are able to experiment with how would-be users react to their products or environments without actually physically building anything. This too allows for significant time and cost savings, especially in the early phases of a design project.

Likewise, other areas of studying human behavior and reactions, such as psychology and psychiatry, have seen similar adoption. By immersing a subject in a virtual environment, practitioners of these fields are able to see how they would react to situations that would otherwise be difficult, if not impossible, to emulate in the real world. Examples of this, discussed in section 2.5, included studying how embodying an avatar that is significantly different from one's self, whether taller (Yee & Bailenson, 2007) or with a different skin color (Peck, Seinfeld, Aglioti & Slater, 2013), affected the behavior of subjects.

In order to be effective, each of these contexts depends on the virtual environment eliciting realistic responses, something that is in turn reliant on maintaining a high sense of presence on the part of the immersed user, as seen in section 2.1. As we have explored in this thesis, maintaining presence in asymmetric virtual reality setups requires the integration of interactions into the virtual environment; if the immersed user perceives interactions as coming from the real world, rather than the virtual one that they are inhabiting, their sense of presence is in danger of being broken or reduced. This requirement for integration was explored in our discussion on presence in asymmetric virtual reality in section 2.3. However, not all integration is equal when it comes to inducing presence. This is where we established our research questions - we were interested in what difference the degree and style of integration had on the immersed user.

To explore these questions, we developed an asymmetric virtual reality environment designed to allow us to test out various styles of integrated interactions. This environment was themed after the concept of escape rooms and featured five puzzles that subjects had to solve in order to escape. Each puzzle involved a different conduit, or method of interaction, in a randomized sequence. These conduits allowed the immersed user to interact and communicate with the external user, the moderator, who provided them with necessary guidance in solving said puzzles. After each session, we had a post-session interview with the subject in order to discern how the interactions had affected their sense of presence and conceptualization of the user.

During our analysis of discussions with users, several themes emerged, as explored in sections 4, 5, and 6. We saw that the degree to which the external user was embodied in the environment had a significant effect on both the immersed user's sense of presence and the dynamic existed between the two.

During higher degrees of external user embodiment, we found that the immersed user's sense of social presence was being affected; the user felt as if they were in the virtual environment with an active participant. During lower degrees, the user's sense of spatial presence was affected instead; spatial cues took the attention off the external user and instead encouraged the user to further immerse themselves in the ambience of the virtual environment, requiring them to navigate the space with increased autonomy and less reliance on external guidance. Here, users often conceptualized the spatial cues that were provided as being gameplay elements rather than hints offered by an actual participant. We also found several interesting middle grounds. One such example was seen with our Voice conduit, during which users conceptualized the external user as a sort of personified gameplay element - the user thought of them as a person, but the interactions were one-way and users would often react to verbal suggestions without acknowledgement, a form of behavior that was significantly different from higher degrees of embodiment, where users would almost always respond with some form of "thank you." Another example of such a middle ground occurred with the Walkie-Talkie conduit, where we saw users conceptualizing the external user as an intermittently present active participant. The external user was indeed seen an active participant in the environment, but only there when the user needed them. This allowed for a balance between focus on the other user and focus on the environment, avoiding the increased feeling of co-presence that came with a higher degree of embodiment.

Whether the user's social presence or spatial presence was being affected, we saw that the reasons for their reductions and breaks in presence remained quite consistent. These reasons were 1) a lack of realism and 2) something in the real world interfering with them. The former, a lack of realism, occurred when the user noticed something that was perceived as being unrealistic. Examples ranged from conduit designs (a common reaction to our Beam of Light conduit) to virtual objects not catching on fire when placed on top of a furnace. The latter reason for breaks in presence, real-world interference, occurred when the user was reminded that they are in virtual reality due to something happening in the actual lab rather than the virtual environment. One of our conduits was intentionally designed to induce this effect, featuring one of the moderators speaking to the user from the real world without any kind of mediation. The virtual reality equipment was another somewhat frequent reason for breaks in presence under this category, with users becoming tangled in the battery cable or having to think about which button to press for a given task. These reasons for breaks in presence were explored in section 2.2.

We also saw several important factors varying based on the degree of external user embodiment. One of these was the degree of control that the external user, or moderator, had over the interactions with the user. When they were highly embodied in the environment, they could easily direct the user's attention and actions using instantaneous verbal prompts. But when they saw a lower degree of embodiment, such as relying on spatial cues to guide the user, they were often left at the whims of the immersed user - waiting for them to notice and react to the guidance. While this does not directly relate to our research questions, it is an important consideration when designing a virtual environment and deciding upon how embodied an external user should be. Other factors, more relevant to dictating how the immersed user perceived the external user's role in the virtual environment, included the degree of social feedback that they received, the degree to which they were aware of the external user, and how much autonomy they had in the virtual environment.

This thesis has been positioned to expand the current literature on presence in virtual reality by looking at how different types of asymmetric interactions affect the immersed user, going past previous studies where the focus was on a single type of integration and whether it was conducive to presence or not. We have shown that depending on the choice of interaction design, there is indeed a significant variance in how the immersed user's sense of presence is affected, as well as how they perceive and interact with the external user. We have also presented the underlying factors that influence this perception, and have offered design suggestions for developing asymmetric virtual reality experiences.

6.1. Limitations

The results of this thesis face several limitations. First, it is difficult to say what the ideal level of external user embodiment is for any given context. Referring back to the examples of asymmetric virtual reality contexts, would a psychiatrist or user experience researcher benefit from a high or low degree of external user embodiment? This is dependent on the dynamic that they want to establish between the two users, requiring insight from practitioners of these fields, making such suggestions extend beyond the scope of this study. Instead, we are simply showing that the degree of embodiment will have significant consequences on the experience and are providing general guidelines to help guide such design decisions. Integrating these findings into specific asymmetric contexts is a promising area of future investigations.

Second, while we can discuss generalized factors such as the degree of social feedback that an immersed user is receiving with any given degree of embodiment, we cannot predict how a user will respond to a specific conduit design. This is clearly illustrated by our Video conduit. In our case, users felt monitored by an authority figure - but that is likely due to the format we chose, a video feed of the external user on the screen. Had it been a virtual avatar instead, reactions could have been very different. Thus, we can say that a high degree of embodiment will lead to the immersed user perceiving the external user as a constantly present active participant in the virtual environment, but whether that participant is a friend or foe depends on the specific design of the conduit being used.

6.2. Managerial implications

As we have shown in this thesis, the degree of external user embodiment has a significant effect on how the immersed user perceives the external user and the dynamic between the two of them. Additionally, factors such as the degree of control that the moderator has during the session were shown to be affected. These findings provide insight into considerations that designers of virtual reality experiences should be mindful of when designing an asymmetric virtual reality experience. Do they want the external user to be a highly present active participant in the environment, and for the immersed user to feel like they are *there* with somebody? Or do they want the external user to be invisible, providing subtle nudges when needed?

The managerial and design implications of each conceptual degree of external user embodiment are succinctly summarized below (Table 4), intended as a reference for virtual environment designers.

Table 4. Degrees of external user embodiment and associated design implications

Degree of external user embodiment	**Unembodied**
Design implications	When the external user is entirely unembodied in the virtual environment, they take on the perceived role of a *gameplay element* rather than an actual person helping the immersed user. Using spatial cues, the external user can subtly nudge the immersed user in the right direction without drawing their attention away from the virtual environment. The immersed user has the highest level of autonomy and control here, with the external user's ability to affect their actions being limited. This degree of embodiment helps prevent the immersed user from feeling monitored or guided.

Degree of external user embodiment	**Un-integrated verbal embodiment**
Design implications	When the external user's embodiment in the virtual environment is limited to only their voice, they take on the perceived role of a *personified gameplay element.* This allows for a neutral, straightforward method of issuing instructions and one-way communications from the external user to the immersed user. Here, the immersed user often treats the external user as part of the gameplay, a narrator-like element rather than an active collaborator, with verbal cues commonly being acted upon without acknowledgement or response. Forming a connection between the users is possible, but not as common as with more embodied methods of interaction.

Degree of external user embodiment	**Integrated verbal embodiment**
Design implications	When the external user is represented in the virtual environment through an interactive physical object, they take on the perceived role of an *intermittently present active participant*. The immersed user considers them to be a person rather than a gameplay element. This degree of embodiment allows for a more sporadic relationship between the two users; while the immersed user conceptualizes the external user as an active participant rather than simply a gameplay element, the external user's presence in the virtual environment is more subtle than with higher degrees of embodiment. They are there if the user needs them, but when interactions are not taking place, the user's attention is freed up from the external participant's presence. This establishes a middle ground between lower degrees of integration, where the user's focus is on the virtual environment, and higher degrees of integration, where the user's focus is on the external user.

Degree of external user embodiment	**Personified verbal embodiment**
Design implications	When the external user is represented in the virtual environment through a character or avatar, sharing the same virtual space, they take on the role of a *constantly present active participant*. This allows for the establishment of a strong sense of connection between the two users, with the immersed user being very aware of the other's presence in the virtual environment. The immersed user receives constant feedback that they are being paid attention to and are being listened to.

We hope that the findings presented in this thesis are useful for practitioners who have either already adopted or are considering adopting virtual reality in their own research.

References

Botella, C., Fernández-Álvarez, J., Guillén, V., García-Palacios, A., & Baños, R. (2017). Recent Progress in Virtual Reality Exposure Therapy for Phobias: A Systematic Review. *Current Psychiatry Reports*, *19*(7). doi: 10.1007/s11920-017-0788-4

Bowman, D., & McMahan, R. (2007). Virtual Reality: How Much Immersion Is Enough?. *Computer*, *40*(7), 36-43. doi: 10.1109/mc.2007.257

Brade, J., Lorenz, M., Busch, M., Hammer, N., Tscheligi, M., & Klimant, P. (2017). Being there again – Presence in real and virtual environments and its relation to usability and user experience using a mobile navigation task. *International Journal Of Human-Computer Studies*, *101*, 76-87. doi: 10.1016/j.ijhcs.2017.01.004

Cairns, P., Cox, A., Day, M., Martin, H., & Perryman, T. (2013). Who but not where: The effect of social play on immersion in digital games. *International Journal Of Human-Computer Studies*, *71*(11), 1069-1077. doi: 10.1016/j.ijhcs.2013.08.015

Cummings, J., & Bailenson, J. (2015). How Immersive Is Enough? A Meta-Analysis of the Effect of Immersive Technology on User Presence. *Media Psychology*, *19*(2), 272-309. doi: 10.1080/15213269.2015.1015740

Diemer, J., Alpers, G., Peperkorn, H., Shiban, Y., & Mühlberger, A. (2015). The impact of perception and presence on emotional reactions: a review of research in virtual reality. *Frontiers In Psychology*, *6*. doi: 10.3389/fpsyg.2015.00026

Frommel, J., Weber, M., Rogers, K., Brich, J., Besserer, D., & Bradatsch, L. et al. (2015). Integrated Questionnaires. *Proceedings Of The 2015 Annual Symposium On Computer-Human Interaction In Play - CHI PLAY '15*. doi: 10.1145/2793107.2793130

Gugenheimer, J., Stemasov, E., Frommel, J., & Rukzio, E. (2017). ShareVR. *Proceedings Of The 2017 CHI Conference On Human Factors In Computing Systems*. doi: 10.1145/3025453.3025683

Han, S., & Kim, J. (2017). A Study on Immersion of Hand Interaction for Mobile Platform Virtual Reality Contents. *Symmetry*, *9*(2), 22. doi: 10.3390/sym9020022

Heeter, C. (1992). Being There: The Subjective Experience of Presence. *Presence: Teleoperators And Virtual Environments*, *1*(2), 262-271. doi: 10.1162/pres.1992.1.2.262

Jeong, K., Kim, J., Kim, M., Lee, J., & Kim, C. (2019). Asymmetric Interface: User Interface of Asymmetric Virtual Reality for New Presence and Experience. *Symmetry*, *12*(1), 53. doi: 10.3390/sym12010053

Kober, S., & Neuper, C. (2013). Personality and Presence in Virtual Reality: Does Their Relationship Depend on the Used Presence Measure?. *International Journal Of Human-Computer Interaction*, *29*(1), 13-25. doi: 10.1080/10447318.2012.668131

Koller, M., Schafer, P., Lochner, D., & Meixner, G. (2019). Rich Interactions in Virtual Reality Exposure Therapy: A Pilot-Study evaluating a System for Presentation

Training. *2019 IEEE International Conference On Healthcare Informatics (ICHI)*. doi: 10.1109/ichi.2019.8904768

Lee, J., Kim, M., & Kim, J. (2019). RoleVR: Multi-experience in immersive virtual reality between co-located HMD and non-HMD users. *Multimedia Tools And Applications*, *79*(1-2), 979-1005. doi: 10.1007/s11042-019-08220-w

Lee, K. (2004). Presence, Explicated. *Communication Theory*, *14*(1), 27-50. doi: 10.1111/j.1468-2885.2004.tb00302.x

Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A Cross-Media Presence Questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators And Virtual Environments*, *10*(3), 282-297. doi: 10.1162/105474601300343612

Ling, Y., Nefs, H., Morina, N., Heynderickx, I., & Brinkman, W. (2014). A Meta-Analysis on the Relationship between Self-Reported Presence and Anxiety in Virtual Reality Exposure Therapy for Anxiety Disorders. *Plos ONE*, *9*(5), e96144. doi: 10.1371/journal.pone.0096144

Liszio, S., & Masuch, M. (2016). Designing Shared Virtual Reality Gaming Experiences in Local Multi-platform Games. *Entertainment Computing - ICEC 2016*, 235-240. doi: 10.1007/978-3-319-46100-7_23

Lombard M, Bolmarcich T, Weinstein L (2009) Measuring Presence: The Temple Presence Inventory

Lombard, M., & Ditton, T. (2006). At the Heart of It All: The Concept of Presence. *Journal Of Computer-Mediated Communication*, *3*(2), 0-0. doi: 10.1111/j.1083-6101.1997.tb00072.x

Mobach, M. (2008). Do virtual worlds create better real worlds?. *Virtual Reality*, *12*(3), 163-179. doi: 10.1007/s10055-008-0081-2

Peck, T., Seinfeld, S., Aglioti, S., & Slater, M. (2013). Putting yourself in the skin of a black avatar reduces implicit racial bias. *Consciousness And Cognition*, *22*(3), 779-787. doi: 10.1016/j.concog.2013.04.016

Price, M., & Anderson, P. (2007). The role of presence in virtual reality exposure therapy. *Journal Of Anxiety Disorders*, *21*(5), 742-751. doi: 10.1016/j.janxdis.2006.11.002

Rebelo, F., Noriega, P., Duarte, E., & Soares, M. (2012). Using Virtual Reality to Assess User Experience. *Human Factors: The Journal Of The Human Factors And Ergonomics Society*, *54*(6), 964-982. doi: 10.1177/0018720812465006

Riches, S., Elghany, S., Garety, P., Rus-Calafell, M., & Valmaggia, L. (2019). Factors Affecting Sense of Presence in a Virtual Reality Social Environment: A Qualitative Study. *Cyberpsychology, Behavior, And Social Networking*, *22*(4), 288-292. doi: 10.1089/cyber.2018.0128

Nieminen, M., & Kirjonen, M. (2020). Keeping It Real!. *Lecture Notes In Computer Science*, 50-60. doi: 10.1007/978-3-030-58465-8_4

Samana, R., Wallach, H., & Safir, M. (2009). The Impact of Personality Traits on the Experience of Presence. *2009 Virtual Rehabilitation International Conference*. doi: 10.1109/icvr.2009.5174197

Schuemie, M., van der Straaten, P., Krijn, M., & van der Mast, C. (2001). Research on Presence in Virtual Reality: A Survey. *Cyberpsychology & Behavior*, *4*(2), 183-201. doi: 10.1089/109493101300117884

Schwind, V., Knierim, P., Chuang, L., & Henze, N. (2017). "Where's Pinky?". *Proceedings Of The Annual Symposium On Computer-Human Interaction In Play*. doi: 10.1145/3116595.3116596

Schwind, V., Knierim, P., Haas, N., & Henze, N. (2019). Using Presence Questionnaires in Virtual Reality. *Proceedings Of The 2019 CHI Conference On Human Factors In Computing Systems - CHI '19*. doi: 10.1145/3290605.3300590

Schwind, V., Knierim, P., Tasci, C., Franczak, P., Haas, N., & Henze, N. (2017). "These are not my hands!". *Proceedings Of The 2017 CHI Conference On Human Factors In Computing Systems*. doi: 10.1145/3025453.3025602

Slater, M. (2003). A note on presence terminology. *Presence Connect*, 3. 1-5.

Slater, M. (2004). How Colorful Was Your Day? Why Questionnaires Cannot Assess Presence in Virtual Environments. *Presence: Teleoperators And Virtual Environments*, *13*(4), 484-493. doi: 10.1162/1054746041944849

Slater, M., Antley, A., Davison, A., Swapp, D., Guger, C., & Barker, C. et al. (2006). A Virtual Reprise of the Stanley Milgram Obedience Experiments. *Plos ONE*, *1*(1), e39. doi: 10.1371/journal.pone.0000039

Slater, M., Khanna, P., Mortensen, J., & Yu, I. (2009). Visual Realism Enhances Realistic Response in an Immersive Virtual Environment. *IEEE Computer Graphics And Applications*, *29*(3), 76-84. doi: 10.1109/mcg.2009.55

Slater, M., & Steed, A. (2000). A Virtual Presence Counter. *Presence: Teleoperators And Virtual Environments*, *9*(5), 413-434. doi: 10.1162/105474600566925

Sugiura, Y., Mochimaru, M., Igarashi, T., Ibayashi, H., Chong, T., & Sakamoto, D. et al. (2018). An asymmetric collaborative system for architectural-scale space design. *Proceedings Of The 16Th ACM SIGGRAPH International Conference On Virtual-Reality Continuum And Its Applications In Industry - VRCAI '18*. doi: 10.1145/3284398.3284416

Tiainen, T., & Jouppila, T. (2019). Use of Virtual Environment and Virtual Prototypes in Co-Design: The Case of Hospital Design. *Computers*, *8*(2), 44. doi: 10.3390/computers8020044

Yee, N., & Bailenson, J. (2007). The Proteus Effect: The Effect of Transformed Self-Representation on Behavior. *Human Communication Research*, *33*(3), 271-290. doi: 10.1111/j.1468-2958.2007.00299.x

Usoh, M., Catena, E., Arman, S., & Slater, M. (2000). Using Presence Questionnaires in Reality. *Presence: Teleoperators And Virtual Environments*, *9*(5), 497-503. doi: 10.1162/105474600566989,

Weech, S., Kenny, S., & Barnett-Cowan, M. (2019). Presence and Cybersickness in Virtual Reality Are Negatively Related: A Review. *Frontiers In Psychology*, *10*. doi: 10.3389/fpsyg.2019.00158

van Baren, J., IJsselsteijn, W. (2004) Measuring Presence: A Guide to Current Measurement Approaches

Zenner, A., Speicher, M., Klingner, S., Degraen, D., Daiber, F., & Krüger, A. (2018). Immersive Notification Framework. *Extended Abstracts Of The 2018 CHI Conference On Human Factors In Computing Systems*. doi: 10.1145/3170427.3188505

Appendices

Appendix One: Participants (anonymized)

Participant ID	Participant year of birth	Participant experience with virtual reality (a lot; some; none)
1	1988	A lot of experience
2	1994	No experience
3	1996	Some experience
4	1987	No experience
5	1991	No experience
6	1992	Some experience
7	1995	No experience
8	1986	Some experience
9	1993	Some experience
10	1994	Some experience
11	1994	No experience
12	1992	Some experience
13	1994	Some experience
14	1994	Some experience
15	1994	Some experience

Appendix Two: Codes

Category: Asymmetric Interactions (factors that are directly related to the research questions)

- Sub-category: Help
  - Code: ROLE OF THE OTHER
  - Code: LOCATION OF THE OTHER Code: AWARENESS OF THE OTHER
- Sub-category: Conduits
  - Code: POSITIVE REACTIONS TO CONDUIT Code: NEGATIVE REACTIONS TO CONDUIT Code: NEUTRAL REACTIONS TO CONDUIT
  - Code: FOCUS ON GAMEPLAY OVER CONDUIT
- Sub-category: Conduit-specific codes
  - Conduit: Spatial conduits (Outlines and Light)
    - Code: UNNOTICED HELP
    - Code: SUBTLE HELP
    - Code: PATRONIZING HELP
    - Code: USEFUL HELP Code: UNREALISTIC
  - Conduit: Yelling
    - Code: BACK TO REALITY
    - Code: PHYSICAL LOCATION OF VOICE
    - Code: DIDN'T MIND
  - Conduit: Voice
    - Code: HELPFUL VOICE
    - Code: PART OF GAMEPLAY
    - Code: CONNECTION
  - Conduit: Walkie talkie
    - Code: ROLEPLAYING
    - Code: ENJOYABLE INTERACTION
    - Code: IGNORING HELP
    - Code: SHARED SPACE
  - Conduit: Video call
    - Code: CONNECTION TO OTHER
    - Code: FEELING WATCHED
    - Code: SHARED SPACE

Category: Gameplay and environment (mediating factors; not directly related but still affect presence)

- Sub-category: Gameplay
  - Code: FOCUS ON GAMEPLAY OVER HELP
  - Code: UNENJOYABLE HELP
  - Code: USEFUL HELP
  - Code: GAMEPLAY RULES
  - Code: COMPARISON TO OTHER EXPERIENCES
  - Code: PERSONALITY FACTOR

Sub-category: Physics interactions
    Code: POSITIVE PHYSICS INTERACTION
    Code: NEGATIVE PHYSICS INTERACTION
    Code: PHYSICS COMMENTS AND QUESTIONS
Sub-category: Visuals and audio
    Code: ENVIRONMENT AND THEME
    Code: QUALITY OF VISUALS
    Code: QUALITY OF AUDIO
    Code: SOURCE OF AUDIO
Sub-category: VR equipment Code: CONTROLLER ISSUES
    Code: HARDWARE ISSUES